

J. Prov. 16/26

UNIVERSIDAD DE GRANADA  
Facultad de Ciencias  
Fecha 22-2-96  
SALIDA NUM. 1270

T  
16  
85

UNIVERSIDAD DE GRANADA  
Facultad de Ciencias  
Fecha 1-12-95  
ENTRADA NUM. 3037

# ALTERNATIVAS GEOMETRICAS EN EL ACP DE UNA V.A. HILBERTIANA

BIBLIOTECA UNIVERSITARIA  
GRANADA  
Nº Documento 019687199  
Nº Copia 121238881

FRANCISCO ANTONIO OCAÑA LARA

Granada, 1995



UNIVERSIDAD DE GRANADA  
27 NOV. 1995  
COMISION DE DOCTORES

D. MARIANO J. VALDERRAMA BONNET, Catedrático de Universidad del área de Estadística e Investigación Operativa en la Universidad de Granada.

Dña. ANA M<sup>a</sup> AGUILERA DEL PINO, Profesora titular interina del área de Estadística e Investigación Operativa en la Universidad de Granada.

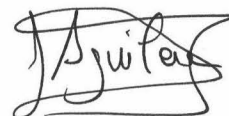
CERTIFICAN:

Que la presente Tesis Doctoral titulada " **Alternativas Geométricas en el ACP de una v.a. hilbertiana** ", ha sido realizada por D. Francisco A. Ocaña Lara, bajo nuestra inmediata dirección, durante el periodo 1993-95, siendo expresión de la capacidad investigadora e interpretativa de su autor, en condiciones que, a nuestro juicio, lo hacen merecedor del título de Doctor en Ciencias, sección Matemáticas.

Granada, 23 de Noviembre de 1995



Prof. Dr. Mariano J. Valderrama Bonnet



Prof. Dra. Ana M<sup>a</sup> Aguilera del Pino

# INTRODUCCION

La descripción de la evolución de una magnitud a lo largo del tiempo es un problema que está presente en un gran número de disciplinas científicas. Históricamente, desde un punto de vista probabilístico y estadístico, su estudio está enmarcado dentro de la Teoría de procesos estocásticos de segundo orden. En dicho marco teórico se han desarrollado un gran número de técnicas, en las que, en la mayoría de los casos, se imponían condiciones que en las aplicaciones resultaban, excesivamente restrictivas, tales como: la estacionariedad, pertenencia a una determinada clase de procesos, (clase de Markov), o incluso, la verificación de un cierto modelo (modelos ARIMA). Si bien es cierto que con dichas técnicas se obtienen excelentes resultados, siempre y cuando el proceso en estudio se encuentre bajo las convenientes condiciones, también es cierto que pueden obtenerse resultados nefastos al intentar aplicar dichas técnicas, fuera de las correspondientes condiciones, ante problemas que afloran en diversas disciplinas científicas y que no se ajustan a esas condiciones.

Una de las técnicas estadísticas, para el análisis de datos funcionales, que menos exigencias requiere es el Análisis en Componentes Principales (ACP). La idea de esta técnica multivariante, extendida por Deville en (1973a) al caso de procesos, es intentar obtener una estructuración de la variabilidad de los datos. Adicionalmente, en el caso de datos funcionales se consigue una descomposición que permite *separar*, por un lado la dependencia con respecto al parámetro continuo (tiempo), y por otro, respecto al propio conjunto de datos.

Esta memoria se estructura en tres capítulos. En el primer capítulo se definen: el ambiente teórico presente a lo largo de toda la memoria, en donde se enmarcan las variables aleatorias hilbertianas, y el ACP de una v.a. hilbertiana. El segundo capítulo está dedicado al problema de la consideración de una estructura geométrica en el *espacio de los datos*. Para finalizar, en el tercer capítulo, se muestran algunos problemas relativos a la modelización mediante componentes principales para el caso de procesos estocásticos.

En el primer capítulo, tras definir el marco teórico en el que consideraremos las v.a. hilbertianas, se define lo que se entiende por el ACP. En su definición, se ha utilizado la proporcionada por Deville (1973a) como generalización del ACP multivariante clásico. Así mismo, desde un punto de vista metodológico se ha hecho especial incapié en el estudio de representaciones ortogonales aso-

cación exitosa del ACP como forma de reducción de la dimensión, en los que el significado de las componentes principales no arroja ninguna información que permita arrojar luz acerca del fenómeno. En especial, en el trabajo de Besse et al. (1986) se van obteniendo diferentes ACP a través de la consideración de diferentes estructuras geométricas sobre un espacio funcional fijo considerado.

En esta memoria se localizan los aspectos relativos a los cambios de estructura geométrica en el espacio de los datos en el segundo capítulo. En particular, en las dos primeras secciones se obtienen los resultados teóricos que nos permitirán resolver e interpretar los casos que se van proponiendo en las siguientes secciones del capítulo.

Por último, el tercer capítulo está dedicado a la modelización de datos funcionales mediante componentes principales. Por un lado, se propone un modelo de predicción mediante componentes principales para un proceso estocástico. Y por otro lado, se aborda el problema de la modelización mediante componentes principales en el caso de disponer de observaciones sobre un conjunto discreto de nodos de observación no necesariamente igualmente espaciados.

Para finalizar, quisiera expresar desde estas líneas, mi profundo agradecimiento al Departamento de Estadística e Investigación Operativa de esta Universidad por su constante estímulo investigador. Así mismo, quisiera agradecer el apoyo material, siempre necesario, prestado por el proyecto PS93-0201, financiado por la DGICYT del Ministerio de Educación y Ciencia. Y por último, y no menos importante, a los directores de esta Tesis, D. Mariano J. Valderrama Bonnet y Dña. Ana M<sup>a</sup> Aguilera del Pino, a los que, a parte de su inestimable ayuda a lo largo de estos años con la dirección científica, quisiera agradecerles sus consejos, su dedicación, tal como si de su propia tesis se tratase, y muy especialmente, el trato personal que me han dispensado.

Granada, Noviembre de 1995

# Indice Temático

<b>1</b>	<b>VARIABLE ALEATORIA HILBERTIANA</b>	<b>3</b>
1.1	INTRODUCCION . . . . .	3
1.2	ESPERANZA MATEMATICA . . . . .	4
1.3	DIAGRAMA DE OPERADORES . . . . .	8
1.3.1	REPRESENTACION DEL OPERADOR COVARIAN- ZA . . . . .	18
1.4	ACP DE UNA V.A. HILBERTIANA . . . . .	22
1.4.1	REPRESENTACIONES ORTOGONALES . . . . .	22
1.4.2	FORMULACION DEL ACP . . . . .	25
1.4.3	ANALISIS ESPECTRAL DE $\mathcal{C}$ y $W$ . . . . .	27
1.4.4	OBTENCION DEL ACP . . . . .	31
1.5	TRANSFORMACION LINEAL DE UNA VARIABLE ALEA- TORIA HILBERTIANA . . . . .	34
1.6	ESTIMACION DEL ACP . . . . .	36
1.6.1	INTRODUCCION . . . . .	36
1.6.2	ACP ESTIMADO . . . . .	43
1.6.3	ALGUNOS RESULTADOS ASINTOTICOS . . . . .	47
<b>2</b>	<b>GEOMETRIAS ALTERNATIVAS EN EL ACP DE UNA V.A. HILBERTIANA</b>	<b>51</b>
2.1	ACP FILTRANTE . . . . .	51
2.1.1	ANALISIS DE LA ESTRUCTURA GEOMETRICA . . . . .	52
2.1.2	ESTUDIO DEL ACP FILTRANTE . . . . .	53
2.1.3	ACP DEL INPUT DE UN SISTEMA ESTOCASTICO LINEAL . . . . .	58
2.2	ESTUDIO SOBRE LA MODIFICACION DE LA GEOMETRIA	66

2.2.1	CONSIDERACIONES SOBRE LA CLASE DE PRODUCTOS ESCALARES EN ESTUDIO . . . . .	66
2.2.2	ACP DE UNA V.A. HILBERTIANA FRENTE A UNA MODIFICACION EN LA GEOMETRIA . . . . .	74
2.3	ACP DE UNA V.A. HILBERTIANA CON VALORES EN UN ESPACIO DE DIMENSION FINITA . . . . .	78
2.3.1	MODIFICACION GEOMETRICA EN EL ESPACIO DE COORDENADAS . . . . .	88
<b>3</b>	<b>MODELIZACION MEDIANTE ACP</b>	<b>93</b>
3.1	PREDICCION LINEAL EN COMPONENTES PRINCIPALES	93
3.1.1	NOTACION Y TEORIA BASICA . . . . .	93
3.1.2	MODELO DE PREDICCION EN COMPONENTES PRINCIPALES . . . . .	96
3.1.3	AJUSTE DE LOS MODELOS DE PREDICCION . . . . .	98
3.2	MODELIZACION CON DATOS DISCRETOS . . . . .	101
3.2.1	COMPARACION DE PREDICCIONES CON DATOS REALES . . . . .	104
3.3	TRATAMIENTO DE NODOS DESIGUALMENTE ESPACIADOS . . . . .	114
<b>A</b>	<b>OBTENCION DE UNA FORMULA DE CUADRATURA</b>	<b>121</b>
A.1	PLANTEAMIENTO DEL PROBLEMA . . . . .	121
A.2	SOBRE EL GRADO DE EXACTITUD DE UNA F. C. . . . .	121
A.3	UN PROBLEMA EQUIVALENTE . . . . .	123
A.4	POLINOMIOS DE LEGENDRE . . . . .	126
A.4.1	Obtención mediante la fórmula de Rodrigues . . . . .	127
A.4.2	Propiedades de los polinomios de Legendre . . . . .	127
A.5	OBTENCION DE LA F. C. GAUSSIANA CON $n + 1$ NODOS	127
A.5.1	Estudio de las integrales $I$ . . . . .	129
A.6	FORMULA DE CUADRATURA GAUSSIANA CON CUATRO NODOS . . . . .	133

# Capítulo 1

## VARIABLE ALEATORIA HILBERTIANA

### §1.1 INTRODUCCION

Este capítulo tiene por objeto definir el ambiente teórico donde trataremos el ACP bajo la perspectiva de una geometría asociada a un cierto espacio de Hilbert. La línea que aquí se seguirá será la de intentar generalizar el marco donde se define el ACP multivariante clásico.

Sean  $(\Omega, \mathcal{A}, P)$  un espacio probabilístico, y  $(H, \langle \cdot, \cdot \rangle_H)$  un espacio de Hilbert separable. En lo que sigue, consideraremos  $X : (\Omega, \mathcal{A}, P) \mapsto H$  una aplicación  $\mathcal{B}_H/\mathcal{A}$  medible.

**Definición 1.1.1** *A una tal aplicación se le denomina **variable aleatoria Hilbert valuada o hilbertiana**. Donde con  $\mathcal{B}_H$  estamos denotando a la  $\sigma$ -álgebra de Borel generada por la topología del espacio  $(H, \langle \cdot, \cdot \rangle_H)$ .*

Antes de comenzar el estudio de las v. a. hilbertianas, objeto de nuestro interés, consideremos el siguiente operador:

**Definición 1.1.2** *Asociado a la variable aleatoria hilbertiana  $X$ , se define el operador*

$$\begin{aligned} u_x : H &\mapsto \mathcal{M}(\Omega, \mathcal{A}) \\ f &\mapsto u_x(f) : (\Omega, \mathcal{A}) \mapsto \mathbb{R} \\ u_x(f)(\omega) &\stackrel{\text{def}}{=} \langle X(\omega), f \rangle_H \quad \forall \omega \in \Omega \end{aligned}$$

Donde  $\mathcal{M}(\Omega, \mathcal{A})$  es el  $\mathbb{R}$ -espacio vectorial formado por las v. a. reales definidas sobre  $(\Omega, \mathcal{A})$

**Lema 1.1.1**  $u_x$  es un operador lineal.

*Demostración*

Sólo probaremos que el operador  $u_x$  está bien definido, puesto que la propiedad de linealidad es inmediata a la vista de la definición.

Dado  $f \in H$ , arbitrario pero fijo, obsérvese que la aplicación  $u_x(f)$  es  $\mathcal{B}_{\mathbb{R}}/\mathcal{A}$  - medible, por poderse obtener mediante la siguiente composición de aplicaciones medibles:

$$\begin{array}{ccc} X & & \langle *, f \rangle_H \\ (\Omega, \mathcal{A}) & \xrightarrow{\text{medible}} & H & \xrightarrow{\text{continua}} & \mathbb{R} \end{array}$$

□

A continuación, nos disponemos a generalizar algunos de los conceptos que mostraron su utilidad en el estudio de vectores aleatorios reales, los cuales, pueden considerarse como un caso particular de v. a. con valores en el espacio de Hilbert (euclídeo)  $\mathbb{R}^n$ .

## §1.2 ESPERANZA MATEMÁTICA

**Definición 1.2.1** Dada  $X$  una v. a. hilbertiana, al elemento de  $H$ , que denotaremos con  $\mathbf{E}[X] \in H$ , verificando

$$\langle \mathbf{E}[X], f \rangle_H = \int_{\Omega} \langle X(\omega), f \rangle_H dP(\omega) \quad \forall f \in H \quad (1.1)$$

se le denomina **esperanza matemática** de dicha v. a. hilbertiana  $X$ .

**Lema 1.2.1** De existir la esperanza matemática de una v.a. hilbertiana, esta es única.

□

A continuación, mostraremos una condición bajo la que puede garantizarse la existencia de la esperanza matemática.



**Proposición 1.2.1** *Supongamos que existe la siguiente integral, que notaremos con*

$$\mathbf{E}[\|X\|_H] \stackrel{\text{def}}{=} \int_{\Omega} \|X(\omega)\|_H dP(\omega) < \infty \quad (1.2)$$

*entonces, existe la esperanza matemática de  $X$ .*

*Demostración*

Comencemos definiendo el siguiente funcional:

$$\Psi_X : (H, \langle \cdot, \cdot \rangle_H) \mapsto \mathbb{R}$$

$$\Psi_X(f) \stackrel{\text{def}}{=} \int_{\Omega} \langle X(\omega), f \rangle_H dP(\omega) \quad \forall f \in H$$

A continuación, estudiaremos el funcional anteriormente definido, el cual cumple las siguientes propiedades:

1.  $\Psi_X$  está bien definido.

Para ello, dado  $f \in H$ , y considerando lo expuesto en el lema (1.1.1), probemos que la v.a. real  $u_x(f)$  es un elemento de  $L_1(\Omega, \mathcal{A}, P)$ . Obsérvese que, aplicando la desigualdad de Cauchy-Schwarz se tiene:

$$\int_{\Omega} |u_x(f)(\omega)| dP(\omega) \leq \|f\|_H \int_{\Omega} \|X(\omega)\|_H dP(\omega) < \infty \quad (1.3)$$

2. Claramente, se observa que  $\Psi_X$  es lineal.
3. Utilizando la expresión anterior, puede concluirse que  $\Psi_X$  es continua, ya que

$$|\Psi_X(f)| \leq \|f\|_H \mathbf{E}[\|X\|_H] \quad \forall f \in H \quad (1.4)$$

Por tanto, siendo  $\Psi_X \in H'$ , puede utilizarse el Teorema de Riesz-Frechet, y asegurar la existencia de un **único** elemento del espacio de hilbert  $H$ , que denotaremos con  $\mathbf{E}[X]$ , verificando

$$\Psi_X(f) = \langle \mathbf{E}[X], f \rangle_H \quad \forall f \in H \quad (1.5)$$

con lo que se tendría garantizado la existencia de la esperanza.

□

A continuación, analizaremos algunas propiedades acerca de la esperanza matemática.

**Proposición 1.2.2** *En las condiciones de la Proposición anterior, se cumple:*

$$\|\mathbf{E}[X]\|_H \leq \mathbf{E}[\|X\|_H]$$

*Demostración*

El Teorema de Riesz-Frechet, aplicado para obtener la ecuación (1.5), establece asimismo que

$$\|\Psi_X\|_{H'} = \|\mathbf{E}[X]\|_H$$

Por otro lado, se concluye a partir de (1.4) que

$$\|\Psi_X\|_{H'} \leq \mathbf{E}[\|X\|_H]$$

Uniendo las dos anteriores expresiones se obtiene lo enunciado en esta Proposición. □

Sin embargo, como muestra el siguiente ejemplo y corrobora el caso particular de que  $H = \mathbb{R}$ , no tendría sentido plantearse si se cumple, en general, la igualdad en la proposición anterior.

**Ejemplo 1.2.1** *Considérese en un espacio probabilístico  $(\Omega, \mathcal{A}, P)$ , siendo  $\mathcal{A}$  una  $\sigma$ -álgebra no trivial, y  $E \in \mathcal{A}$  un suceso tal que  $0 < P(E) < 1$ .*

*Con estas premisas, y dado  $f \in H - \{0\}$ , definimos la variable aleatoria hilbertiana*

$$X(w) \stackrel{\text{def}}{=} \begin{cases} f & \text{si } w \in E \\ -f & \text{si } w \in E^c \end{cases} \quad \forall w \in \Omega$$

Puede comprobarse que dicha variable cumple (1.2), obteniéndose además que

$$\mathbf{E}[\|X\|_H] = \|f\|_H$$

Por otro lado, aplicando la definición de esperanza se obtiene que

$$\mathbf{E}[X] = (2P(E) - 1)f \implies \|\mathbf{E}[X]\|_H = |2P(E) - 1| \|f\|_H$$

Y teniendo en cuenta lo anteriormente obtenido, se concluye que

$$\|\mathbf{E}[X]\|_H < \mathbf{E}[\|X\|_H]$$

sin más que tener presente que  $|2P(E) - 1| < 1$ .

□

A continuación, consideraremos el  $\mathbb{R}$ -espacio de Banach

$$L_1(\Omega, \mathcal{A}, P; H) \stackrel{\text{def}}{=} \left\{ X : (\Omega, \mathcal{A}, P) \mapsto H : \begin{array}{l} X \text{ es } \mathcal{B}_H/\mathcal{A} \text{ - medible} \\ \int_{\Omega} \|X(\omega)\|_H dP(\omega) < \infty \end{array} \right\}$$

sobre el cuál, según lo establecido en la Proposición (1.2.1), tendremos definido lo que se denomina **operador esperanza**, dado por

$$E : L_1(\Omega, \mathcal{A}, P; H) \mapsto (H, \langle *, * \rangle_H)$$

**Proposición 1.2.3** *El operador esperanza  $E$  es lineal y continuo.*

*Demostración*

En principio, es evidente, a partir de la definición de esperanza, que dicho operador es lineal.

En relación a la continuidad, utilizando la Proposición (1.2.2), puede observarse que  $\forall X \in L_1(\Omega, \mathcal{A}, P; H)$  se cumple:

$$\|\mathbf{E}[X]\|_H \leq \|X\|_{L_1(\Omega, \mathcal{A}, P; H)}$$

□

**Proposición 1.2.4** *La restricción del operador esperanza  $E$  al espacio de Hilbert  $L_2(\Omega, \mathcal{A}, P; H)$  puede interpretarse como una proyección ortogonal sobre el subespacio vectorial de  $L_2(\Omega, \mathcal{A}, P; H)$*

$$\{f I_{\Omega} : f \in H\} \subseteq L_2(\Omega, \mathcal{A}, P; H)$$

*isométricamente isomorfo a  $H$ .*

*Demostración*

Haciendo abuso de notación, denotaremos con  $E$  a la aplicación

$$E : H \longrightarrow \{f I_\Omega : f \in H\}$$

$$E(X) = \mathbf{E}[X] I_\Omega \quad \forall X \in L_2(\Omega, \mathcal{A}, P; H)$$

Dado  $X \in L_2(\Omega, \mathcal{A}, P; H)$ , arbitrario pero fijo, probaremos que  $E(X)$  resulta ser la proyección ortogonal de  $X$  sobre el subespacio vectorial (cerrado) considerado. Para ello, denotando con  $\langle *, * \rangle$  al producto escalar de  $L_2(\Omega, \mathcal{A}, P; H)$ , puede comprobarse que  $\forall f \in H$

$$\langle X, f I_\Omega \rangle = \int_\Omega \langle X(\omega), f \rangle_H = \langle \mathbf{E}[X], f \rangle_H = \langle E(X), f I_\Omega \rangle$$

□

### §1.3 DIAGRAMA DE OPERADORES

El punto de inicio de esta sección consiste en generalizar el concepto de *matriz de covarianzas* al caso que nos ocupa. En este ambiente mostraremos cómo, asociada a una v.a. hilbertiana, podemos considerar un diagrama de operadores, en los cuales aparecerán descrita toda la información, tanto probabilística como geométrica, del ACP de la v.a. considerada.

Para todas las v.a. hilbertianas, que se consideren en esta sección, se supondrá la existencia de su *esperanza matemática*.

**Definición 1.3.1** Dada  $X$  una v.a. hilbertiana, se define su **operador de covarianza asociado**, como el operador

$$C_x : H \longmapsto H$$

definido por la condición:  $\forall f, g \in H$

$$\langle C_x(f), g \rangle_H = \int_\Omega \langle X(\omega) - \mathbf{E}[X], f \rangle_H \langle X(\omega) - \mathbf{E}[X], g \rangle_H dP(\omega) \quad (1.1)$$

**Lema 1.3.1** De existir el operador de covarianza, asociado a una v.a. hilbertiana, este es único.

□

A continuación, enunciaremos algunas propiedades que son inmediatas de la definición.

**Proposición 1.3.1** *El operador de covarianza  $C_x$  es un operador lineal, simétrico y semidefinido positivo.*

□

Como en el caso de la esperanza, a continuación mostramos una condición bajo la que se demostrará la existencia del operador de covarianza.

A partir de ahora, supondremos que  $X$  es una v.a. hilbertiana cumpliendo

$$\mathbf{E} [\|X\|_H^2] \stackrel{\text{def}}{=} \int_{\Omega} \|X(\omega)\|_H^2 dP(\omega) < \infty \quad (1.2)$$

Bajo la anterior premisa, más adelante se mostrará la existencia y se estudiarán algunas propiedades que cumple dicho operador de covarianza. Antes de continuar con el estudio del operador de covarianza, se definen unos operadores que proporcionarán el marco ideal para tratar, entre otros, a dicho operador de covarianza.

En principio, por ser  $X$  una v.a. hilbertiana cumpliendo (1.2), puede asegurarse la existencia de la esperanza,  $\mathbf{E}[X] \in H$ . A dicha esperanza, por comodidad en la notación, la denotaremos con:

$$\mu \stackrel{\text{def}}{=} \mathbf{E}[X]$$

**Definición 1.3.2** *Asociado a la variable aleatoria hilbertiana  $X$ , se define el operador*

$$\begin{aligned} U_x : H &\longmapsto L_2(\Omega, \mathcal{A}, P) \\ f &\longmapsto U_x(f) : (\Omega, \mathcal{A}, P) \longmapsto \mathbb{R} \\ U_x(f)(\omega) &\stackrel{\text{def}}{=} \langle X(\omega) - \mu, f \rangle_H \quad \forall \omega \in \Omega \end{aligned}$$

De ahora en adelante, denotaremos con  $\tilde{X}$  a la v.a. hilbertiana dada por

$$\tilde{X}(\omega) \stackrel{\text{def}}{=} X(\omega) - \mu \quad \forall \omega \in \Omega$$

A continuación, enumeraremos algunas propiedades del operador  $U_x$  definido anteriormente.

**Proposición 1.3.2**  $U_x$  es un operador lineal y continuo, cumpliéndose además

$$\|U_x\|_B \leq \|\tilde{X}\|_{L_2(\Omega, \mathcal{A}, P; H)}$$

denotando con  $\|\cdot\|_B$  a la norma del espacio  $BL(H, L_2(\Omega, \mathcal{A}, P))$ .

*Demostración*

1. El operador  $U_x$  está bien definido.

En principio, dado  $f \in H$  arbitrario pero fijo, por lo demostrado en el Lema (1.1.1) la aplicación

$$U_x(f) = u_{\tilde{x}}(f) \text{ es } \mathcal{B}_{\mathbb{R}}/\mathcal{A} \text{ - medible.}$$

Así mismo, probemos que  $U_x(f) \in L_2(\Omega, \mathcal{A}, P)$

$$\int_{\Omega} U_x(f)^2(w) dP(w) \leq \|f\|_H^2 \int_{\Omega} \|\tilde{X}(w)\|_H^2 dP(w) < \infty \quad (1.3)$$

2.  $U_x$  es lineal.
3.  $U_x$  es continuo, es decir,  $U_x \in BL(H, L_2(\Omega, \mathcal{A}, P))$

Haciendo uso de la expresión (1.3) se tendría

$$\|U_x(f)\|_2^2 \leq \|f\|_H^2 \mathbf{E} \left[ \|\tilde{X}\|_H^2 \right] < \infty \quad \forall f \in H$$

obteniéndose, de esta forma que:

$$\|U_x\|_B \leq \|\tilde{X}\|_{L_2(\Omega, \mathcal{A}, P; H)}$$

Donde estamos denotando con  $\|\cdot\|_2$  a la norma de  $L_2(\Omega, \mathcal{A}, P)$ .

□

**Proposición 1.3.3** La aplicación

$$U : L_2(\Omega, \mathcal{A}, P; H) \mapsto BL(H, L_2(\Omega, \mathcal{A}, P))$$

es lineal y continua.

*Demostración*

En principio, la linealidad de la aplicación  $U$  es obvia, a partir de la definición.

Por otro lado, obsérvese que de la proposición anterior se obtiene que para cada  $X \in L_2(\Omega, \mathcal{A}, P; H)$

$$\begin{aligned} \|U_x\|_B^2 &\leq \|\tilde{X}\|_{L_2(\Omega, \mathcal{A}, P; H)}^2 = \mathbf{E} [\|X - \mu\|_H^2] \leq \\ &\leq 2 \left( \mathbf{E} [\|X\|_H^2] + \|\mu\|_H^2 \right) \leq 2 \left( \mathbf{E} [\|X\|_H^2] + \mathbf{E} [\|X\|_H^2]^2 \right) \\ &\leq 2 \left( \mathbf{E} [\|X\|_H^2] + \mathbf{E} [\|X\|_H^2] \right) = 4 \|X\|_{L_2(\Omega, \mathcal{A}, P; H)}^2 \end{aligned}$$

□

**Proposición 1.3.4**  $U_x$  es además un operador compacto y de Hilbert-Schmidt.

*Demostración*

1.  $U_x$  es un operador compacto (completamente continuo).

Para demostrar esta propiedad, consideraremos una sucesión de v.a. hilbertianas centradas,  $\{Z_n : n \in \mathbb{N}\} \in L_2(\Omega, \mathcal{A}, P; H)$ , tales que

- convergen a  $\tilde{X}$  en el espacio  $L_2(\Omega, \mathcal{A}, P; H)$ .
- cada v.a.  $Z_n$  es de la forma siguiente

$$Z_n = \sum_{i=1}^{\sigma(n)} f_{i,n} I_{A(i,n)}$$

siendo  $\sigma(n) \in \mathbb{N}$ ,  $f_{i,n} \in H$ , e  $I_{A(i,n)}$  resulta ser la función indicadora del suceso  $A(i, n) \in \mathcal{A}$

En estas condiciones, asociada a cada v.a.  $Z_n$ , podemos considerar su operador asociado  $U_n \stackrel{\text{def}}{=} U_{Z_n}$ . Dicho operador puede comprobarse que es un operador lineal de rango finito, en concreto, para cada  $n \in \mathbb{N}$

$$U_n(H) \subseteq \text{Lin} \{I_{A(i,n)} : i = 1, \dots, \sigma(n)\}$$

De la Proposición anterior, se concluye que el operador  $U_x$  resulta ser el límite de la anterior sucesión de operadores lineales de rango finito en el espacio de Banach  $BL(H, L_2(\Omega, \mathcal{A}, P))$ . Y por tanto,  $U_x$  resulta ser un operador compacto.

2.  $U_x$  es un operador de Hilbert-Schmidt.

Sea  $\mathcal{B} = \{e_i : i \in I\}$  una base ortonormal del espacio de Hilbert  $H$ , siendo  $I$  un conjunto numerable. Y consideremos la *serie* en estudio

$$\sum_{i \in I} \mathbf{E} [U_x(e_i)^2] = \sum_{i \in I} \int_{\Omega} \langle \tilde{X}(w), e_i \rangle_H^2 dP(w)$$

Por ser  $\mathcal{B}$  una base ortonormal de  $H$ , para cada  $w \in \Omega$ , se cumple lo siguiente:

$$\begin{aligned} \|\tilde{X}(w)\|_H^2 &= \sum_{i \in I} \langle \tilde{X}(w), e_i \rangle_H^2 = \\ &= \sum_{i \in I} U_x(e_i)^2(w) \quad (\text{convergencia en } \mathbb{R}) \end{aligned}$$

Es decir, la serie  $\sum_{i \in I} U_x(e_i)^2$  es una serie que converge puntualmente a  $\|\tilde{X}\|_H^2$ , y está formada por v.a. reales no negativas e integrables. Así mismo, por la hipótesis (1.2), la sucesión de sumas parciales estaría dominada por la v.a.  $\|\tilde{X}\|_H^2 \in L_1(\Omega, \mathcal{A}, P)$ . Por tanto, aplicando el Teorema de la Convergencia Dominada se tiene:

$$\mathbf{E} \left[ \|\tilde{X}\|_H^2 \right] = \mathbf{E} \left[ \sum_{i \in I} U_x(e_i)^2 \right] = \sum_{i \in I} \mathbf{E} [U_x(e_i)^2]$$

Concluimos de lo anterior que  $U_x$  es un operador de Hilbert-Schmidt, y además, que su norma en dicho espacio resulta ser:

$$\|U_x\|_{HS}^2 = \mathbf{E} \left[ \|\tilde{X}\|_H^2 \right]$$

□



**Corolario 1.3.1** *La aplicación*

$$U : L_2(\Omega, \mathcal{A}, P; H) \mapsto HS(H, L_2(\Omega, \mathcal{A}, P))$$

*es lineal y continua.*

*Demostración*

Teniendo en cuenta la última expresión obtenida en la demostración anterior, y por lo demostrado en la Proposición (1.3.3) se obtiene:

$$\|U_x\|_{HS}^2 = \mathbf{E} \left[ \|\tilde{X}\|_H^2 \right] \leq 4 \|X\|_{L_2(\Omega, \mathcal{A}, P; H)}^2$$

□

**Definición 1.3.3** *En las condiciones en las que se definió el operador  $U_x$ , puede garantizarse la existencia de su operador adjunto, que denotaremos con*

$$\begin{aligned} U_x^* : L_2(\Omega, \mathcal{A}, P) &\mapsto H \\ \psi &\mapsto U_x^*(\psi) \end{aligned}$$

*y definido por la siguiente relación*

$$\langle U_x^*(\psi), f \rangle_H = \langle \psi, U_x(f) \rangle_2 \quad \forall \psi \in L_2(\Omega, \mathcal{A}, P) \quad \forall f \in H$$

*Donde estamos denotando con  $\langle *, * \rangle_2$  al producto escalar del espacio de Hilbert  $L_2(\Omega, \mathcal{A}, P)$ .*

**Proposición 1.3.5** *En las condiciones anteriores, se tiene*

$$U_x^*(\psi) = \mathbf{E}[\psi(X - \mu)] \quad \forall \psi \in L_2(\Omega, \mathcal{A}, P)$$

*Donde con  $\psi(X - \mu)$  estamos denotando a la variable aleatoria de  $L_1(\Omega, \mathcal{A}, P; H)$  definida por*

$$(\psi(X - \mu))(w) \stackrel{\text{def}}{=} \psi(w)(X(w) - \mu) \quad \forall w \in \Omega$$

*Demostración*

Obsérvese que dado  $\psi \in L_2(\Omega, \mathcal{A}, P)$ , se tiene que  $\forall f \in H$  se cumple

$$\langle \psi, U_x(f) \rangle_2 = \int_{\Omega} \psi(w) U_x(f)(w) dP(w) =$$

$$= \int_{\Omega} \langle \psi(w) (X(w) - \mu), f \rangle_H dP(w) = \langle U_x^*(\psi), f \rangle_H$$

Obsérvese que si la aplicación

$$w \in \Omega \longrightarrow \psi(w) (X(w) - \mu)$$

fuese una variable aleatoria hilbertiana de  $L_1(\Omega, \mathcal{A}, P; H)$ , por la anterior expresión se concluiría que

$$U_x^*(\psi) = \mathbf{E}[\psi (X - \mu)]$$

con lo que tendríamos demostrado esta Proposición.

A continuación probaremos que

$$\psi (X - \mu) \in L_1(\Omega, \mathcal{A}, P; H)$$

1.  $\psi (X - \mu)$  es  $\mathcal{B}_H/\mathcal{A}$  medible, por poderse obtener como composición de aplicaciones medibles.
- 2.

$$\int_{\Omega} \|\psi(w) (X(w) - \mu)\|_H dP(w) \leq \|\psi\|_2 \left( \mathbf{E} [\|X - \mu\|_H^2] \right)^{\frac{1}{2}} < \infty$$

□

En la siguiente proposición se enumeran algunas propiedades que cumple  $U_x^*$ , por el hecho de ser el operador adjunto de  $U_x$ .

**Proposición 1.3.6**  $U_x^*$  es un operador lineal, continuo, compacto, Hilbert-Schmidt, y además cumple:

$$\|U_x^*\| = \|U_x\|$$

donde con  $\|\cdot\|$  estamos denotando a cualquiera de las normas consideradas en este caso. Es decir, tanto la del correspondiente espacio de operadores lineales y continuos, como la del espacio de los operadores de Hilbert-Schmidt.

**Proposición 1.3.7** *La aplicación*

$$U^* : L_2(\Omega, \mathcal{A}, P; H) \mapsto BL(L_2(\Omega, \mathcal{A}, P), H)$$

*es lineal y continua.*

*Además, el resultado sigue siendo cierto si, en lugar de considerar el espacio de Banach  $BL(L_2(\Omega, \mathcal{A}, P), H)$ , se considera el espacio de Hilbert formado por los operadores de Hilbert-Schmidt de dicho espacio de Banach.*

**Proposición 1.3.8**

$$C_x = U_x^* \circ U_x$$

*Demostración*

Obsérvese que dado  $\forall f, g \in H$  se tiene

$$\begin{aligned} \langle (U_x^* \circ U_x)(f), g \rangle_H &= \langle U_x(f), U_x(g) \rangle_2 = \\ &= \int_{\Omega} \langle X(w) - \mu, f \rangle_H \langle X(w) - \mu, g \rangle_H dP(w) \end{aligned}$$

Por tanto, se puede concluir de lo anterior que

$$C_x(f) = (U_x^* \circ U_x)(f) \quad \forall f \in H$$

□

**Definición 1.3.4** *Dada  $X \in L_2(\Omega, \mathcal{A}, P; H)$ , se define el operador de Escoufier asociado, al obtenido de la forma siguiente*

$$W_x \stackrel{\text{def}}{=} U_x \circ U_x^*$$

Obsérvese que el operador de Escoufier es de la forma siguiente

$$\begin{aligned} W_x : L_2(\Omega, \mathcal{A}, P) &\mapsto L_2(\Omega, \mathcal{A}, P) \\ \psi &\mapsto W_x(\psi) : (\Omega, \mathcal{A}, P) \mapsto \mathbb{R} \end{aligned}$$

$$W_x(\psi)(w) \stackrel{\text{def}}{=} \int_{\Omega} \langle X(w) - \mu, X(\sigma) - \mu \rangle_H \psi(\sigma) dP(\sigma) \quad \forall w \in \Omega$$

Puesto que  $\forall w \in \Omega$  se tiene que

$$\begin{aligned}
W_x(\psi)(w) &= U_x(U_x^*(\psi))(w) = U_x(\mathbf{E}[\psi(X - \mu)])(w) = \\
&= \langle X(w) - \mu, \mathbf{E}[\psi(X - \mu)] \rangle_H = \\
&= \int_{\Omega} \langle \psi(\sigma)(X(\sigma) - \mu), X(w) - \mu \rangle_H dP(\sigma)
\end{aligned}$$

**Proposición 1.3.9** *De la propia definición, se deduce que el operador de Escoufier es un operador lineal, simétrico y semidefinido positivo.*

**Teorema 1.3.1** *Ambos operadores  $C_x$  y  $W_x$  son: lineales, simétricos, semidefinidos positivos, continuos, compactos, Hilbert-Schmidt, y nucleares.*

*Demostración*

En primer lugar, las propiedades de linealidad, continuidad y compacidad de los operadores mencionados vienen dadas a partir de las mismas propiedades correspondientes a los operadores  $U_x$  y  $U_x^*$ .

La demostración de las restantes propiedades, para los operadores de covarianza y de Escoufier, siguen un mismo esquema. Por esta razón, demostraremos sólo las propiedades para el operador de covarianza.

En lo que sigue, denotaremos con  $\mathcal{B} = \{e_i : i \in I\}$  a una base ortonormal de  $H$ .

En relación a la pertenencia al espacio de los operadores nucleares, puede observarse que se cumple

$$\begin{aligned}
\sum_{i \in I} \langle C_x(e_i), e_i \rangle_H &= \sum_{i \in I} \langle U_x^* \circ U_x(e_i), e_i \rangle_H = \\
&= \sum_{i \in I} \|U_x(e_i)\|_2^2 = \|U_x\|_{HS}^2 < \infty
\end{aligned}$$

por ser  $U_x$  un operador de Hilbert-Schmidt.

Por otro lado, para analizar si el operador de  $C_x$  es de Hilbert-Schmidt estudiemos la serie  $\sum_{i \in I} \|C_x(e_i)\|_H^2$

Obsérvese que por ser  $U_x^*$  un operador lineal continuo se tiene, para cada  $i \in I$ , la desigualdad

$$\|C_x(e_i)\|_H = \|U_x^*(U_x(e_i))\|_H \leq \|U_x^*\| \|U_x(e_i)\|_2$$

donde con  $\|U_x^*\|$  estamos denotando a la norma del espacio  $BL(L_2(\Omega, \mathcal{A}, P), H)$ .

Por tanto, haciendo uso de esta desigualdad se obtiene

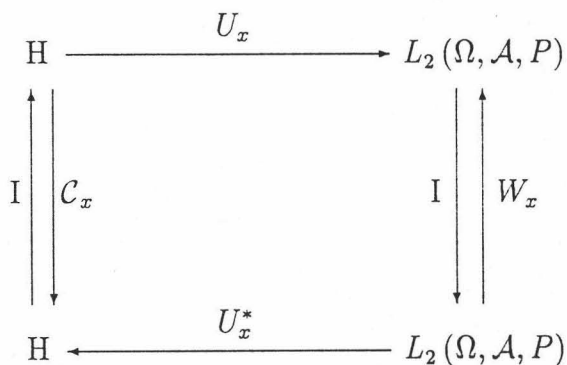
$$\sum_{i \in I} \|C_x(e_i)\|_H^2 \leq \|U_x^*\|^2 \sum_{i \in I} \|U_x(e_i)\|_2^2 < \infty$$

con lo que se deduce que  $C_x$  es un operador de Hilbert-Schmidt.

□

**Corolario 1.3.2**  $tr(C_x) = tr(W_x) = \mathbf{E} \left[ \|\tilde{X}\|_H^2 \right]$

Todo lo anterior puede recopilarse en el siguiente diagrama de operadores, que recoge las estructuras, hilbertiana y probabilística, de la v.a. hilbertiana considerada



### 1.3.1 REPRESENTACION DEL OPERADOR COVARIANZA

Sea  $X : (\Omega, \mathcal{A}, P) \mapsto H$  una v.a. de  $L_2(\Omega, \mathcal{A}, P; H)$  para la que supondremos, por comodidad en la notación, que  $\mathbf{E}[X] = 0$ . En estas condiciones, podemos considerar el operador de covarianza asociado,  $C_x : H \mapsto H$ , definido por la propiedad siguiente:

$$\langle C_x(f), g \rangle_H = \mathbf{E}[\langle X, f \rangle_H \langle X, g \rangle_H] \quad \forall f, g \in H$$

Pretendemos probar, entre otros resultados, que el operador de covarianza puede expresarse de la forma siguiente:

$$C_x = \mathbf{E}[X \otimes X]$$

expresión que generaliza la definición de matriz de covarianza en el caso de vectores aleatorios reales.

**Lema 1.3.2** *Para cualesquiera  $\{f, g\} \subseteq H$ , el operador  $f \otimes g : H \mapsto H$  definido por*

$$f \otimes g(x) = \langle x, f \rangle_H g \quad \forall x \in H$$

*es un operador lineal, continuo y de Hilbert-Schmidt siendo sus normas, en ambos espacios, coincidentes*

$$\|f \otimes g\|_{HS} = \|f \otimes g\|_{BL(H)} = \|f\|_H \|g\|_H$$

*Demostración*

En principio, la propiedad de linealidad es inmediata de la propia definición. Para la comprobación de la continuidad, obsérvese que

$$\begin{aligned} \|f \otimes g\|_{BL(H)} &\stackrel{\text{def}}{=} \sup_{x \in \mathcal{S}_H} \|f \otimes g(x)\|_H = \\ &= \|g\|_H \sup_{x \in \mathcal{S}_H} |\langle x, f \rangle_H| = \|g\|_H \|f\|_H \end{aligned}$$

Por otra parte, dada una base ortonormal de  $H$ ,  $\mathcal{B} = \{e_i : i \in I\}$ , consideremos la serie siguiente:

$$\begin{aligned} \|f \otimes g\|_{HS}^2 &\stackrel{\text{def}}{=} \sum_{i \in I} \|f \otimes g(e_i)\|_H^2 = \\ &= \sum_{i \in I} \langle e_i, f \rangle_H^2 \|g\|_H^2 = \|g\|_H^2 \|f\|_H^2 \end{aligned}$$

□

**Lema 1.3.3** *La aplicación  $L : H \mapsto HS$ ,  $L(f) \stackrel{\text{def}}{=} f \otimes f \quad \forall f \in H$ , es continua.*

*Observación: Este resultado también sería cierto si, en lugar de considerar el espacio de los operadores Hilbert-Schmidt (HS), se considerase el espacio  $BL(H)$ .*

*Demostración*

Dada una sucesión  $\{f_n\}_{n \in \mathbb{N}} \subseteq H$  convergente a  $f \in H$ , es decir,

$$\lim_{n \rightarrow \infty} \|f_n - f\|_H = 0,$$

probaremos que la sucesión  $\{L(f_n)\}_{n \in \mathbb{N}} \subseteq HS$  converge en  $HS$  a  $L(f)$ .

En primer lugar, dado  $n \in \mathbb{N}$  y  $x \in H$ , se tiene

$$\begin{aligned} (L(f) - L(f_n))(x) &= \langle x, f \rangle_H f - \langle x, f_n \rangle_H f_n = \\ &= \langle x, f \rangle_H (f - f_n) - \langle x, f_n - f \rangle_H f_n \end{aligned}$$

de donde se obtiene que

$$L(f) - L(f_n) = f \otimes (f - f_n) - (f_n - f) \otimes f_n$$

Aplicando la norma de Hilbert-Schmidt a la igualdad de operadores anterior, y utilizando el Lema anterior, obtenemos lo siguiente:

$$\|L(f) - L(f_n)\|_{HS} \leq \|f\|_H \|f - f_n\|_H + \|f_n - f\|_H \|f_n\|_H$$

de donde se concluye  $\lim_{n \rightarrow \infty} \|L(f) - L(f_n)\|_{HS} = 0$

□

**Lema 1.3.4** *Sean  $G : H \mapsto H$  un operador de Hilbert-Schmidt, y  $\mathcal{B} = \{e_i : i \in I\}$  una base ortonormal de  $H$ .*

*Entonces, para cualesquiera  $l, j \in I$*

$$\langle G, e_l \otimes e_j \rangle_{HS} = \langle G(e_l), e_j \rangle_H$$

*Demostración*

$$\begin{aligned} \langle G, e_l \otimes e_k \rangle_{HS} &= \sum_{i \in I} \langle G(e_i), e_l \otimes e_j(e_i) \rangle_H = \\ &= \sum_{i \in I} \langle e_l, e_i \rangle_H \langle G(e_i), e_j \rangle_H = \langle G(e_l), e_j \rangle_H \end{aligned}$$

□

**Teorema 1.3.2** *El operador covarianza es un operador de Hilbert-Schmidt, que puede obtenerse como sigue*

$$C_x = \mathbf{E}[X \otimes X]$$

*Demostración*

En primer lugar, por lo visto en el Lema (1.3.2), la aplicación  $X \otimes X : (\Omega, \mathcal{A}, P) \rightarrow HS$  dada por

$$X \otimes X(w) \stackrel{\text{def}}{=} X(w) \otimes X(w) \quad \forall w \in \Omega$$

está bien definida.

En relación a la medibilidad, obsérvese que dicha aplicación podría obtenerse mediante la siguiente composición de aplicaciones medibles

$$\begin{array}{ccccc} & X & & L & \\ (\Omega, \mathcal{A}, P) & \xrightarrow{\text{medible}} & H & \xrightarrow{\text{continua}} & HS \end{array}$$

con lo cuál, dicha aplicación es  $\mathcal{B}_{HS}/\mathcal{A}$ -medible.

Por otro lado, utilizando el Lema (1.3.2) se obtiene

$$\int_{\Omega} \|X \otimes X(w)\|_{HS}^2 dP(w) = \int_{\Omega} \|X(w)\|_H^2 dP(w) < \infty$$

con lo que podemos asegurar la existencia de  $\mathbf{E}[X \otimes X] \in HS$  caracterizado por la propiedad siguiente:

$$\langle \mathbf{E}[X \otimes X], \Gamma \rangle_{HS} = \int_{\Omega} \langle X \otimes X(w), \Gamma \rangle_{HS} dP(w) \quad \forall \Gamma \in HS$$



A continuación, probaremos que el operador de Hilbert-Schmidt  $\mathbf{E}[X \otimes X]$  resulta ser el operador de covarianza de  $X$ . Para ello, consideramos una base  $\mathcal{B} = \{e_i : i \in I\}$  ortonormal de  $H$ , con la que haciendo uso del Lema (1.3.4), obtenemos que  $\forall j, l \in I$

$$\begin{aligned} \langle \mathbf{E}[X \otimes X](e_l), e_j \rangle_H &= \langle \mathbf{E}[X \otimes X], e_l \otimes e_j \rangle_{HS} = \\ \int_{\Omega} \langle X \otimes X(w), e_l \otimes e_j \rangle_{HS} &= \int_{\Omega} \langle X(w) \otimes X(w)(e_l), e_j \rangle_H = \\ \int_{\Omega} \langle X(w), e_l \rangle_H \langle X(w), e_j \rangle_H &= \langle C_x(e_l), e_j \rangle_H \end{aligned}$$

Concluyendo así que  $\mathbf{E}[X \otimes X] = C_x$

□

**Proposición 1.3.10** *Para cada  $f \in H$ , se obtiene que*

$$C_x(f) = \mathbf{E}[X \otimes X(f)]$$

siendo  $X \otimes X(f)$  la v.a. de  $L_1(\Omega, \mathcal{A}, P; H)$  dada por

$$X \otimes X(f)(w) = \langle X(w), f \rangle_H X(w) \quad \forall f \in H$$

*Demostración*

La clave de este resultado está en la siguiente igualdad:  $\forall f, g \in H$

$$\int_{\Omega} \langle X(w), f \rangle_H \langle X(w), g \rangle_H = \int_{\Omega} \langle X(w) \otimes X(w)(f), g \rangle_H$$

De esta forma, de ser la aplicación medible con valores en  $H$

$$w \rightarrow X(w) \otimes X(w)(f) = \langle X(w), f \rangle_H X(w)$$

de  $L_1(\Omega, \mathcal{A}, P; H)$ , podría también concluirse, para cualesquiera  $f, g \in H$ , que

$$\langle \mathbf{E}[X \otimes X(f)], g \rangle_H = \int_{\Omega} \langle X(w) \otimes X(w)(f), g \rangle_H = \langle C_x(f), g \rangle_H$$

y por tanto que  $C_x(f) = \mathbf{E}[X \otimes X(f)] \quad \forall f \in H$

Para terminar, la comprobación de que  $X \otimes X(f) \in L_1(\Omega, \mathcal{A}, P; H)$  viene dada por

$$\int_{\Omega} \|X(w) \otimes X(w)(f)\|_H dP(w) \leq \|f\|_H \int_{\Omega} \|X(w)\|_H^2 dP(w) < \infty$$

□

## §1.4 ACP DE UNA V.A. HILBERTIANA

A lo largo de esta sección, consideramos  $X : (\Omega, \mathcal{A}, P) \mapsto H$  una v.a. hilbertiana de  $L_2(\Omega, \mathcal{A}, P; H)$  que, para simplificar la notación, supondremos centrada ( $\mathbf{E}[X] = 0$ ).

### 1.4.1 REPRESENTACIONES ORTOGONALES

En principio, por tomar valores  $X$  en el espacio  $H$ , para cualquier base ortonormal del espacio de hilbert  $(H, \langle *, * \rangle_H)$ ,  $B = \{e_i : i \in I\}$  con  $I \subseteq \mathbb{N}$ , podemos construir una representación, para cada  $w \in \Omega$ , de la forma siguiente:

$$X(w) = \sum_{i \in I} \eta_i(w) e_i \quad (1.1)$$

donde la *suma* anterior converge en  $(H, \langle *, * \rangle_H)$ , y donde estamos notando con

$$\eta_i(w) \stackrel{\text{def}}{=} \langle X(w), e_i \rangle_H = U_x(e_i)(w) \quad \forall w \in \Omega \quad \forall i \in I$$

**Proposición 1.4.1** *Teniendo en cuenta la descomposición anterior (1.1), se verifica:*

1.  $\eta_i \in L_2(\Omega, \mathcal{A}, P) \quad \forall i \in I$
2.  $\mathbf{E}[\eta_i] = 0 \quad \forall i \in I$
3.  $\sum_{i \in I} \text{Var}[\eta_i] = \mathbf{E}[\|X\|_H^2] = \text{tr}(\mathcal{C})$

#### *Demostración*

En principio, estas propiedades son consecuencia inmediata de los resultados obtenidos al estudiar el operador  $U_x$ .

□

Como es sabido, la representación (1.1), en cada  $w \in \Omega$ , es convergente en  $(H, \langle *, * \rangle_H)$ , es decir, se trata de una convergencia *puntual*. A continuación, intentaremos garantizar también la convergencia de dicha representación en el espacio  $L_2(\Omega, \mathcal{A}, P; H)$ .

**Lema 1.4.1** Sean  $\xi : (\Omega, \mathcal{A}) \mapsto (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$  una variable aleatoria  $\mathcal{B}_{\mathbb{R}}/\mathcal{A}$  medible, y  $f \in H$ . En esta situación la aplicación  $Z : \Omega \mapsto H$

$$Z(w) \stackrel{\text{def}}{=} \xi(w) f \quad \forall w \in \Omega$$

es  $\mathcal{B}_H/\mathcal{A}$  medible.

Además, si la variable  $\xi \in L_2(\Omega, \mathcal{A}, P)$  se cumple que  $Z \in L_2(\Omega, \mathcal{A}, P; H)$ .

*Demostración*

Para demostrar este resultado, analicemos en primer lugar la siguiente aplicación

$$\begin{aligned} L : \mathbb{R} &\mapsto H \\ L(r) &\stackrel{\text{def}}{=} r f \quad \forall r \in \mathbb{R} \end{aligned}$$

la cuál, sin dificultad se demuestra, que es lineal y continua.

Teniendo en cuenta lo anteriormente comentado, obsérvese que la aplicación  $Z$  podría obtenerse mediante la siguiente composición de aplicaciones medibles:

$$\begin{array}{ccccc} & \xi & & L & \\ (\Omega, \mathcal{A}, P) & \xrightarrow{\text{medible}} & \mathbb{R} & \xrightarrow{\text{continua}} & H \end{array}$$

Además, en el caso de que  $\xi \in L_2(\Omega, \mathcal{A}, P)$  se tiene

$$\int_{\Omega} \|Z(w)\|_H^2 dP(w) = \|f\|_H^2 \mathbf{E}[\xi^2] < \infty$$

□

**Proposición 1.4.2** La descomposición (1.1) converge en  $L_2(\Omega, \mathcal{A}, P; H)$ .

*Demostración*

En primer lugar, dado  $J \subseteq I$  un subconjunto finito, haciendo uso del lema anterior, puede afirmarse que

$$\sum_{j \in J} \eta_j e_j \in L_2(\Omega, \mathcal{A}, P; H)$$

Además, notando con  $\|\cdot\|$  a la norma del espacio  $L_2(\Omega, \mathcal{A}, P; H)$  se razona como sigue:

$$\begin{aligned} \left\| X - \sum_{j \in J} \eta_j e_j \right\|^2 &= \int_{\Omega} \left\| X(w) - \sum_{j \in J} \eta_j e_j(w) \right\|_H^2 dP(w) = \\ & \int_{\Omega} \|X(w)\|_H^2 dP(w) + \sum_{j \in J} \mathbf{E} [\eta_j^2] - 2 \int_{\Omega} \langle X(w), \sum_{j \in J} \eta_j(w) e_j \rangle_H dP(w) \end{aligned}$$

Obsérvese que el último sumando de la expresión anterior quedaría de la forma siguiente:

$$\sum_{j \in J} \int_{\Omega} \eta_j(w) \langle X(w), e_j \rangle_H dP(w) = \sum_{j \in J} \mathbf{E} [\eta_j^2]$$

En definitiva, se tendría que

$$\left\| X - \sum_{j \in J} \eta_j e_j \right\|^2 = \mathbf{E} [\|X\|_H^2] - \sum_{j \in J} \mathbf{E} [\eta_j^2]$$

y haciendo uso de lo establecido en la propiedad 3 de la Proposición (1.4.1) se obtiene la convergencia, de la serie en cuestión, en  $L_2(\Omega, \mathcal{A}, P; H)$ .

□

**Corolario 1.4.1** *La descomposición (1.1) converge en probabilidad, es decir, para cada  $\epsilon > 0$  se cumple*

$$\lim_{n \rightarrow \infty} P \left[ \left\| X - \sum_{j \in J_n} \eta_j e_j \right\|_H > \epsilon \right] = 0$$

donde  $J_n = \{i \in I : i \leq n\}$

1.4.2 FORMULACION DEL ACP

En la sección anterior se ha obtenido que, dada una v.a. hilbertiana de segundo orden centrada, cualquier base ortonormal del espacio  $(H, \langle *, * \rangle_H)$  permite construir una representación en términos de, además de los elemento de dicha base, una colección v.a. centradas reales de  $L_2(\Omega, \mathcal{A}, P)$ .

Ahora, se pretende analizar cómo obtener la base ortonormal del espacio  $H$ , que proporcione una representación *óptima*, en un sentido que más adelante se especificará. Dicha propiedad de optimalidad, vendrá definida a partir de unas propiedades a exigir sobre las v.a. reales, a obtener en la correspondiente representación.

En definitiva, el esquema que se seguirá aquí resulta ser una generalización del propuesto por Hötelling para el caso de v.a. con valores en  $\mathbb{R}^n$ .

**Definición 1.4.1** Asociada a la v.a. hilbertiana centrada  $X \in L_2(\Omega, \mathcal{A}, P; H)$ , se denominan componentes principales a una sucesión de v.a. reales  $\{\psi_i : i \in I\}$  con  $I = \{1, 2, \dots\} \subseteq \mathbb{N}$ , verificando:

- $\{\psi_i : i \in I\} \subseteq U_x(S_H)$
- $\mathbf{E}[\psi_i] = 0 \quad \forall i \in I$  (Obsérvese que esta propiedad es redundante.)

•

$$\text{Var}[\psi_1] = \max \{ \text{Var}[\eta] : \eta \in U_x(S_H) \}$$

Si  $i > 1$

$$\text{Var}[\psi_i] = \max \left\{ \text{Var}[\eta] \middle/ \begin{array}{l} \eta \in U_x(S_H) \\ \mathbf{E}[\eta \psi_j] = 0 \quad \forall j < i \end{array} \right\}$$

**Proposición 1.4.3** Las v.a. reales de  $U_x(H)$  verifican:

1.  $U_x(H) \subseteq L_2(\Omega, \mathcal{A}, P)$
2.  $\mathbf{E}[U_x(f)] = 0 \quad \forall f \in H$
3.  $\text{cov}(U_x(f), U_x(g)) = \langle C_x(f), g \rangle_H \quad \forall f, g \in H$
4.  $\text{Var}[U_x(f)] = \langle C_x(f), f \rangle_H \quad \forall f \in H$

*Demostración*

La demostración de este apartado es consecuencia inmediata de las propiedades del operador  $U_x$ .

En primer lugar, obsérvese que para cada  $f \in H$ , por ser  $\mathbf{E}[X] = 0$  se cumple:

$$\mathbf{E}[U_x(f)] = \mathbf{E}[\langle X, f \rangle_H] = \langle \mathbf{E}[X], f \rangle_H = 0$$

Por otro lado, dados  $f, g \in H$ , se cumple:

$$\begin{aligned} \text{cov}(U_x(f), U_x(g)) &= \mathbf{E}[U_x(f)U_x(g)] = \langle U_x(f), U_x(g) \rangle_2 = \\ &= \langle U_x^* \circ U_x(f), g \rangle_H = \langle C_x(f), g \rangle_H \end{aligned}$$

□

De lo anterior, puede afirmarse que el problema de encontrar la colección de componentes principales, resulta ser equivalente a encontrar  $\{f_i : i \in I\} \subseteq S_H$  cumpliendo:

$$\langle C_x(f_1), f_1 \rangle_H = \max \{ \langle C_x(f), f \rangle_H / f \in S_H \}$$

Si  $i > 1$

$$\langle C_x(f_i), f_i \rangle_H = \max \left\{ \langle C_x(f), f \rangle_H / \begin{array}{l} f \in S_H \\ \langle C_x(f), f_j \rangle_H = 0 \quad \forall j < i \end{array} \right\}$$

Existe un resultado, obtenido por Hilbert y cuya demostración puede verse en (Riesz-Nagy, 1990, pp. 230), que establece que el anterior problema, planteado en términos del operador  $C_x$ , tiene como solución una base ortonormal de  $H$  formada por vectores propios  $\{f_i : i \in I\}$ , que consideraremos asociado a la sucesión decreciente  $(\lambda_i : i \in I)$ , en la que aparecen los valores propios repetidos tantas veces como indique su multiplicidad, es decir,

$$\lambda_1 \geq \lambda_2 \geq \dots \geq 0$$

$$C_x(f_i) = \lambda_i f_i \quad \forall i \in I$$

**Definición 1.4.2** *A los elementos de una base ortonormal,  $\{f_i : i \in I$ , que determine un conjunto de componentes principales, se les denomina factores principales.*

### 1.4.3 ANALISIS ESPECTRAL DE $C$ y $W$

A continuación, pretendemos analizar algunas propiedades relativas al espectro de los operadores de covarianza y de Escoufier. El punto de partida es el Teorema (1.3.1), en donde, se enumeran las propiedades comunes de dichos operadores, y a partir del cuál, utilizaremos la Teoría Espectral de operadores lineales simétricos y compactos sobre un espacio de Hilbert.

En nuestro estudio, para referirnos simultáneamente a los operadores de covarianza y Escoufier, en calidad sólo de operadores que cumplen las propiedades anteriormente citadas del Teorema (1.3.1), consideraremos el operador

$$F : E \mapsto E$$

lineal, simétrico, semidefinido positivo y compacto (continuo), definido sobre el espacio de Hilbert  $E$ .

En el análisis espectral de operadores sobre espacios vectoriales con dimensión infinita, el concepto de espectro es más amplio que en el caso finito dimensional. En concreto, haciendo uso de la notación introducida, puede hablarse de:

- Espectro puntual  $P\sigma(F)$

$$\lambda \in P\sigma(F) \stackrel{\text{def}}{\iff} \text{Ker}(F - \lambda I) \neq \{0\}$$

- Espectro continuo  $C\sigma(F)$

$$\lambda \in C\sigma(F) \stackrel{\text{def}}{\iff} (F - \lambda I)^{-1} \text{ existe sobre un conjunto denso de } E \text{ y no es continua.}$$

En nuestro caso, se cumple que  $\sigma(F) = P\sigma(F) \cup C\sigma(F)$

A pesar de esta distinción, en la descomposición espectral del operador  $F$  jugará un papel fundamental su espectro puntual.

**Proposición 1.4.4** *El operador  $F$  cumple las siguientes propiedades:*

1.  $\sigma(F) = P\sigma(F) \cup \{0\}$
2.  $\sigma(F) - \{0\} = P\sigma(F) - \{0\}$

3. Si  $\lambda \in P\sigma(F) - \{0\}$  el subespacio vectorial  $\text{Ker}(F - \lambda I)$  es de dimensión finita.
4.  $\sigma(F)$  es numerable y acotado (compacto).
5.  $\sigma(F) \subseteq \mathbb{R}_0^*$
6.  $\sigma(F)' = \{0\}$
7. Todos los subespacios propios son ortogonales entre sí.

□

**Teorema 1.4.1**

$$\sigma(C) - \{0\} = \sigma(W) - \{0\}$$

Además, para cualquier valor propio  $\lambda$ , del anterior conjunto de valores propios, se tiene la siguiente relación entre los correspondientes subespacios propios:

$$E_\lambda = U^*(E_\lambda^*) \quad E_\lambda^* = U(E_\lambda)$$

donde  $E_\lambda \stackrel{\text{def}}{=} \text{Ker}(C - \lambda I)$  y  $E_\lambda^* \stackrel{\text{def}}{=} \text{Ker}(W - \lambda I)$

*Demostración*

En principio, dado  $\lambda \in P\sigma(C) - \{0\}$ , ha de existir entonces  $f \in H - \{0\}$  verificando

$$C(f) = \lambda f$$

Si aplicamos el operador  $U$  a ambos miembros de la expresión anterior se obtiene

$$U \circ C(f) = W(U(f)) = \lambda U(f)$$

Con el objeto de poder concluir que  $\lambda \in P\sigma(W)$ , comprobaremos que  $U(f) \neq 0$ . Para ello, obsérvese que

$$\|U(f)\|_2^2 = \langle U(f), U(f) \rangle_2 = \langle f, C(f) \rangle_H = \lambda \|f\|_H^2 \neq 0$$

De esta forma, podremos concluir que

$$P\sigma(C) - \{0\} \subseteq P\sigma(W) - \{0\}$$

$$U(E_\lambda) \subseteq E_\lambda^*$$



y aplicando el operador  $U^*$  a la anterior expresión se obtendría

$$U^* \circ U(E_\lambda) = C(E_\lambda) = E_\lambda \subseteq U^*(E_\lambda^*)$$

Repitiendo un razonamiento similar al anterior para el operador  $W$ , concluimos que:

$$\begin{aligned} P\sigma(W) - \{0\} &\subseteq P\sigma(C) - \{0\} \\ U^*(E_\lambda^*) &\subseteq E_\lambda \end{aligned}$$

y ahora aplicando el operador  $U$  a la última expresión se obtiene

$$U \circ U^*(E_\lambda^*) = W(E_\lambda^*) = E_\lambda^* \subseteq U(E_\lambda)$$

□

**Corolario 1.4.2** *Si consideráramos el conjunto*

$$P\sigma(F) - \{0\} = \{\lambda_k : k \in K\}$$

donde  $K = \{1, 2, \dots\} \subseteq \mathbb{N}$ , estando ordenado como sigue:

$$\lambda_n > \lambda_m \quad \forall n < m \in K$$

*En el caso de que el conjunto  $P\sigma(F)$  sea infinito se cumple:*

$$\lim_{k \rightarrow \infty} \lambda_k = 0$$

*Demostración*

Por lo expuesto en la Proposición anterior, la sucesión  $\{\lambda_k : k \in K\}$  sería una sucesión decreciente y acotada en  $\mathbb{R}$ , formada por elementos distintos entre sí. Con lo cuál, dicha sucesión será convergente hacia el único elemento de  $\sigma(F)' = \{0\}$ .

□

**Teorema 1.4.2** *En estas condiciones, existe una base ortonormal de  $E$  formada por vectores propios del operador  $F$ . Cumpliéndose además, que el operador  $F$  admite la siguiente representación*

$$F = \sum_{\lambda \in P\sigma(F)^*} \lambda \pi_\lambda \tag{1.2}$$

en el espacio de Banach  $BL(E)$ . Donde estamos denotando con

$$P\sigma(F)^* \stackrel{\text{def}}{=} P\sigma(F) - \{0\}$$

$\pi_\lambda$  proyección ortogonal en el subespacio propio  $\text{Ker}(F - \lambda I)$

En determinadas condiciones, además de la convergencia anterior, se cumple:

1. Si  $F$  es un operador de Hilbert-Schmidt, entonces, la representación (1.2) converge en el espacio de los operadores de Hilbert-Schmidt ( $HS(E)$ ).
2. Si  $F$  es un operador nuclear, entonces, dicha representación (1.2) converge en el espacio de los operadores nucleares ( $N(E)$ ).

*Demostración*

En principio, la convergencia de la descomposición (1.2) viene dada en el Teorema de Descomposición Espectral de los operadores lineales simétricos y compactos.

En segundo lugar, consideremos una base ortonormal  $\mathcal{B} = \{e_i : i \in I\}$  de  $H$  formada por vectores propios del operador  $F$ . Y así mismo, denotemos con  $\lambda_i \in P\sigma(F)$  al valor propio asociado al vector  $e_i$ , y con  $I = \{1, 2, \dots\} \subseteq \mathbb{N}$ . De forma que en la sucesión  $(\lambda_i : i \in I)$  se repite cada valor propio, tantas veces como indique su multiplicidad.

Por otro lado, obsérvese que en estas condiciones se tiene:

$$\sum_{\lambda \in P\sigma(F)^*} \lambda \pi_\lambda = \sum_{i \in I} \lambda_i e_i \otimes e_i$$

Consideremos, en las siguientes demostraciones,  $J_n \subseteq I$  un subconjunto finito, tal que  $J_n \stackrel{\text{def}}{=} \{i \in I : i \leq n\}$

1. Supongamos que  $F$  es de Hilbert-Schmidt

$$\begin{aligned} \left\| F - \sum_{j \in J_n} \lambda_j e_j \otimes e_j \right\|_{HS(E)}^2 &= \sum_{i \in I} \left\| F(e_i) - \sum_{j \in J_n} \lambda_j e_j \otimes e_j(e_i) \right\|_H^2 = \\ &= \sum_{i > n} \|F(e_i)\|_H^2 \rightarrow 0 \text{ cuando } n \rightarrow \infty \end{aligned}$$

2. Supongamos que  $F$  es nuclear

$$\begin{aligned} \left\| F - \sum_{j \in J_n} \lambda_j e_j \otimes e_j \right\|_{N(E)} &= \sum_{i \in I} \langle F(e_i) - \sum_{j \in J_n} \lambda_j e_j \otimes e_j(e_i), e_i \rangle_H = \\ &= \sum_{i > n} \langle F(e_i), e_i \rangle_H \rightarrow 0 \text{ cuando } n \rightarrow \infty \end{aligned}$$

□

**Proposición 1.4.5** Si  $F$  es un operador Hilbert-Schmidt y nuclear se obtiene:

$$\|F\|_{BL(E)} = \sup_{i \in I} \lambda_i$$

$$\|F\|_{HS(E)} = \left( \sum_{i \in I} \lambda_i^2 \right)^{\frac{1}{2}}$$

$$\|F\|_{N(E)} = \sum_{i \in I} \lambda_i$$

siendo  $(\lambda_i : i \in I)$  la sucesión de valores propios, en la que aparece cada elemento de  $P\sigma(F)$  repetido, tantas veces como indica su multiplicidad.

#### 1.4.4 OBTENCION DEL ACP

Consideramos  $\mathcal{B} = \{f_i : i \in I\}$  una base ortonormal de  $H$  formada por vectores propios del operador  $C_x$ , es decir

$$C_x(f_i) = \lambda_i f_i$$

De forma que la sucesión  $(\lambda_i : i \in I)$  en la que aparecen los valores propios, repetidos tantas veces como indica su multiplicidad, está ordenada como sigue:

$$\lambda_i > \lambda_j \geq 0 \quad \forall i < j$$

**Teorema 1.4.3** Cualquier base ortonormal  $\mathcal{B}$  del espacio  $H$ , donde diagonalice  $C_x$ , determina una colección de componentes principales, de la forma siguiente:

$$\{U_x(f_i) : i \in I\}$$

*Demostración*

Este teorema fue obtenido por Hilbert, y una demostración puede verse en (Riesz-Nagy, 1990, pp. 230).

□

**OBSERVACIONES**

- Las componentes principales obtenidas a partir de elementos del  $KerC_x$ , son v.a. reales nulas. Es decir, si  $C_x(f_i) = 0$  la componente principal asociada  $U_x(f_i)$  verificaría

$$E[U_x(f_i)] = 0 \quad Var[U_x(f_i)] = 0$$

con lo cuál  $U_x(f_i) = 0$  (c.s).

De esta forma, en la correspondiente representación ortogonal de la v.a. hilbertiana en término de la colección de componentes principales, las asociadas al  $KerC_x$  podría ser obviadas.

Por esta razón, en lugar de considerar una base ortonormal de vectores propios de  $C_x$  del espacio  $H$ , la consideraremos del espacio  $(KerC_x)^{\perp H}$ . Con lo que en definitiva estaremos considerando la colección de vectores propios asociados a valores propios no nulos  $P\sigma(C_x) - \{0\}$

- Puede observarse que podrían existir situaciones en las que existiesen. para una misma v.a. hilbertiana, distintas colecciones de componentes principales.
- Así mismo, a parte de la representación ortogonal que de la v.a. hilbertiana permiten las componentes principales. Dicha representación tiene asociadas sendas representaciones para los operadores de covarianza y Escoufier

$$C_x = \sum_{i \in I^*} \lambda_i f_i \otimes f_i$$

$$W_x = \sum_{i \in I^*} \xi_i \otimes \xi_i$$

la cuál converge en los correspondientes espacios de los operadores lineales y continuos, Hilbert-Schmidt y nucleares. Y donde estamos notando con

$$I^* = \{i \in I : \lambda_i > 0\}$$

cumpléndose además:

1.  $\xi_i = U_x(f_i) \quad \forall i \in I$
2.  $W_x(\xi_i) = \lambda_i \xi_i \quad \forall i \in I$
3.  $U_x^*(\xi_i) = \lambda_i f_i \quad \forall i \in I$
4.  $\mathbf{E}[\xi_i \xi_j] = \lambda_i \delta(i, j) \quad \forall i, j \in I$

### §1.5 TRANSFORMACION LINEAL DE UNA VARIABLE ALEATORIA HILBERTIANA

Sean  $X : (\Omega, \mathcal{A}, P) \mapsto H$  una v.a. hilbertiana, y  $F : H \mapsto G$  un operador lineal y continuo ( $F \in BL(H, G)$ ) con valores en el espacio de Hilbert  $G$ . En esta sección, denotaremos con  $\|\cdot\|_B$  a la asociada al espacio de Banach  $BL(H, G)$ .

En principio, por ser  $F$  un operador continuo, podemos considerar la v.a. hilbertiana siguiente

$$Y \stackrel{\text{def}}{=} F \circ X : (\Omega, \mathcal{A}, P) \mapsto G$$

El objeto de esta sección es analizar: las propiedades de  $Y$  heredadas de la v.a.  $X$ , así como la relación existentes entre sus correspondientes operadores asociados.

**Proposición 1.5.1** Si  $X \in L_1(\Omega, \mathcal{A}, P; H)$  entonces  $Y \in L_1(\Omega, \mathcal{A}, P; G)$ . Cumpliéndose además que  $\mathbf{E}[Y] = F(\mathbf{E}[X])$ .

*Demostración*

En principio, obsérvese que se cumple:

$$\mathbf{E}[\|Y\|_G] \stackrel{\text{def}}{=} \int_{\Omega} \|F(X(w))\|_G dP(w) \leq \|F\|_B \mathbf{E}[\|X\|_H] < \infty$$

con lo que  $Y \in L_1(\Omega, \mathcal{A}, P; G)$ .

Por otro lado, dado  $g \in G$  arbitrario pero fijo, se cumple que

$$\begin{aligned} \langle \mathbf{E}[Y], g \rangle_G &= \int_{\Omega} \langle F(X(w)), g \rangle_G dP(w) = \\ &= \int_{\Omega} \langle X(w), F^*(g) \rangle_H = \langle \mathbf{E}[X], F^*(g) \rangle_H = \langle F(\mathbf{E}[X]), g \rangle_G \end{aligned}$$

de donde se concluye que  $\mathbf{E}[Y] = F(\mathbf{E}[X])$ .

□

**Proposición 1.5.2** Si  $X \in L_2(\Omega, \mathcal{A}, P; H)$  entonces  $Y \in L_2(\Omega, \mathcal{A}, P; G)$ . Además, se cumple lo siguiente:

1. Si denotamos con  $\tilde{Y}(w) \stackrel{\text{def}}{=} Y(w) - \mathbf{E}[Y]$  y con  $\tilde{X}(w) \stackrel{\text{def}}{=} X(w) - \mathbf{E}[X]$  se obtiene:

$$\tilde{Y}(w) = F(\tilde{X}(w))$$

2.  $U_y = U_x \circ F^*$

3.  $U_y^* = F \circ U_x^*$

4.  $C_y = F \circ C_x \circ F^*$

5.  $W_y = U_x \circ F^* \circ F \circ U_x^*$

*Demostración*

En principio, obsérvese que se cumple:

$$\mathbf{E}[\|Y\|_G^2] \stackrel{\text{def}}{=} \int_{\Omega} \|F(X(w))\|_G^2 dP(w) \leq \|F\|_B^2 \mathbf{E}[\|X\|_H^2] < \infty$$

con lo que  $Y \in L_2(\Omega, \mathcal{A}, P; G)$ .

Por otra parte, haciendo uso de la Proposición anterior y de la linealidad de  $F$ , se cumple, para cada  $w \in \Omega$ , lo siguiente:

$$\tilde{Y}(w) = F(X(w)) - F(\mathbf{E}[X]) = F(\tilde{X}(w))$$

En esta situación, sabemos que podemos definir cuatro operadores asociados a esta v.a. hilbertiana  $Y$  de segundo orden, los cuales, entre otras cosas, nos permiten definir el ACP. Lo que pretendemos analizar es la relación existente entre dichos operadores y los correspondientes a la v.a.  $X$ . Obsérvese que una vez demostrado el primer punto, los demás son consecuencia inmediata de este y de las definiciones de los correspondientes operadores.

Sabemos que  $U_y : G \mapsto L_2(\Omega, \mathcal{A}, P)$  es un operador definido, para cada  $g \in G$ , de la forma siguiente:

$$\begin{aligned} U_y(g)(w) &\stackrel{\text{def}}{=} \langle \tilde{Y}(w), g \rangle_G = \langle F(\tilde{X}(w)), g \rangle_G = \\ &= \langle \tilde{X}(w), F^*(g) \rangle_H = U_x(F^*(g))(w) \quad \forall w \in \Omega \end{aligned}$$

con lo cuál

$$U_y(g) = U_x(F^*(g)) = U_x \circ F^*(g) \quad \forall g \in G$$

□



## §1.6 ESTIMACION DEL ACP

### 1.6.1 INTRODUCCION

Consideremos un espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ , y un espacio de Hilbert  $(H, \langle \cdot, \cdot \rangle_H)$ . En este ambiente, dada una v.a. hilbertiana

$$X : (\Omega, \mathcal{A}, P) \mapsto H \in L_2(\Omega, \mathcal{A}, P; H),$$

un ACP suyo vendría dado a partir de los siguientes elementos:

- $\mu \stackrel{\text{def}}{=} \mathbf{E}[X] \in H$
- $C_X \stackrel{\text{def}}{=} \mathbf{E}[\tilde{X} \otimes \tilde{X}] : H \rightarrow H \in HS$   
siendo  $\tilde{X} \stackrel{\text{def}}{=} X - \mu$ .

donde estamos denotando con  $HS$  al espacio de Hilbert de los operadores de Hilbert-Schmidt del espacio  $H$  en sí mismo.

Supóngase que se consideran  $N$  realizaciones *independientes* del experimento aleatorio asociado a  $X$ , obteniendo así,  $N$  observaciones de la v.a.  $X$ . es decir

$$\{X(w_i^*) : i = 1, \dots, N\} \subseteq H$$

Esta situación puede modelarse considerando una muestra aleatoria simple (m.a.s.) de  $X$  de tamaño  $N$ , que denotaremos con

$$\{X^{(i)} : i = 1, \dots, N\}$$

donde, para cada  $i \in \{1, \dots, N\}$ , estamos considerando

$$X^{(i)} : (\Omega^N, \mathcal{A}^N, P^N) \mapsto (H, \mathcal{B}_H)$$

$$X^{(i)}(w) \stackrel{\text{def}}{=} X(w_i) \quad \forall w = (w_1, \dots, w_N) \in \Omega^N$$

siendo  $(\Omega^N, \mathcal{A}^N, P^N)$  el espacio probabilístico producto *definido en cada componente* a partir del espacio  $(\Omega, \mathcal{A}, P)$ .

En estas condiciones, el conjunto de las  $N$ -observaciones obtenidas para la v.a.  $X$ , podrían considerarse como *una realización* de la anterior m.a.s.

$$\{X^{(i)}(w^*) : i = 1, \dots, N\} \subseteq H \quad \text{siendo } w^* = (w_1^*, \dots, w_N^*).$$



Obsérvese que el mismo esquema teórico serviría para el caso de que se considerase  $N = \infty$ .

**Proposición 1.6.1** *La familia de aplicaciones  $\{X^{(i)} : i = 1, \dots, N\}$  cumple:*

1. *Todas son  $\mathcal{B}_H/\mathcal{A}^N$  -medibles.*
2. *Todas tienen la misma distribución, e igual a su vez a la distribución de la v.a.  $X$ .*
3. *Son independientes entre sí.*
4. *Todas son elementos de  $L_2(\Omega^N, \mathcal{A}^N, P^N; H)$  cumpliendo, para cada  $i \in \{1, \dots, N\}$ , lo siguiente:*

$$\mathbf{E}[X^{(i)}] = \mu \quad \mathcal{C}_{X^{(i)}} \stackrel{\text{def}}{=} \mathbf{E}[\tilde{X}^{(i)} \otimes \tilde{X}^{(i)}] = \mathcal{C}_X$$

denotamos con  $\tilde{X}^{(i)} \stackrel{\text{def}}{=} X^{(i)} - \mu$ .

*Demostración*

En primer lugar, para demostrar la medibilidad, obsérvese que cada v. a.  $X^{(i)}$  se obtiene mediante la siguiente composición de aplicaciones medibles

$$\begin{array}{ccccc} & \pi_i & & X & \\ (\Omega^N, \mathcal{A}^N) & \xrightarrow{\text{medible}} & (\Omega, \mathcal{A}) & \xrightarrow{\text{medible}} & (H, \mathcal{B}_H) \end{array}$$

siendo:  $\pi_i(w) \stackrel{\text{def}}{=} w_i \quad \forall w = (w_1, \dots, w_N) \in \Omega^N$ .

En relación a la distribución de probabilidad de cada v. a. hilbertiana  $X^{(i)}$ , se cumple que para cada  $E \in \mathcal{B}_H$

$$P^N[X^{(i)} \in E] = P^N[\Omega^{N-1} \times [X \in E]] = P[X \in E]$$

Por otra parte, en relación a la independencia estocástica, considérense un conjunto finito de índices,  $J \subseteq \{1, \dots, N\}$  ( con  $\text{card}(J) \leq N$ ), y una familia  $\{E_j : j \in J\} \subseteq \mathcal{B}_H$ . En estas condiciones se verifica:

$$\begin{aligned}
P^N [X^{(j)} \in E_j / \forall j \in J] &= P^N \left[ \prod_{j \in J} [X \in E_j] \times \Omega^{N-\text{card}(J)} \right] = \\
&= \prod_{j \in J} P [X \in E_j] = \prod_{j \in J} P^N [X^{(j)} \in E_j]
\end{aligned}$$

Por último, sin más que hacer uso del Teorema de Fubini, y utilizando las definiciones de la media y del operador de covarianza, se tendría el último punto de la Proposición. Concretamente, para la v.a.  $X^{(i)}$  se razonaría como sigue:

- Para cada  $f \in H$  se cumple

$$\begin{aligned}
\int_{\Omega^N} \langle X^{(i)}(w), f \rangle_H dP^N(w) &= \int_{\Omega^N} \langle X(w_i), f \rangle_H dP^N(w) = \\
&= \int_{\Omega} \langle X(w_i), f \rangle_H dP(w_i) = \langle \mu, f \rangle_H
\end{aligned}$$

- Para cualesquiera  $f, g \in H$ , se cumple

$$\begin{aligned}
\langle \mathcal{C}_{X^{(i)}}(f), g \rangle_H &= \int_{\Omega^N} \langle \tilde{X}^{(i)}(w), f \rangle_H \langle \tilde{X}^{(i)}(w), g \rangle_H dP^N(w) = \\
&= \int_{\Omega} \langle \tilde{X}(w_i), f \rangle_H \langle \tilde{X}(w_i), g \rangle_H dP(w_i) = \langle \mathcal{C}_X(f), g \rangle_H
\end{aligned}$$

□

Obsérvese cómo el esquema anterior determina una m.a.s. de la v.a. hilbertiana  $\tilde{X} \otimes \tilde{X} \in L_1(\Omega, \mathcal{A}, P; HS)$  de tamaño  $N$  dada por

$$\{\tilde{X}^{(i)} \otimes \tilde{X}^{(i)} : i = 1, \dots, N\} \subseteq L_1(\Omega^N, \mathcal{A}^N, P^N; HS)$$

donde, para cada  $i \in \{1, \dots, N\}$ ,  $\tilde{X}^{(i)} \otimes \tilde{X}^{(i)} : (\Omega^N, \mathcal{A}^N, P^N) \mapsto (HS, \mathcal{B}_{HS})$  resulta ser el  $i$ -ésimo elemento de dicha m.a.s., es decir:

$$\begin{aligned}
\tilde{X}^{(i)} \otimes \tilde{X}^{(i)}(w) &\stackrel{\text{def}}{=} \tilde{X}^{(i)}(w) \otimes \tilde{X}^{(i)}(w) = \\
&= \tilde{X}(w_i) \otimes \tilde{X}(w_i) = \tilde{X} \otimes \tilde{X}(w_i) \quad \forall w = (w_1, \dots, w_N) \in \Omega^N
\end{aligned}$$

Concretamente, podríamos enunciar resultados similares a los enunciados en la Proposición (1.6.1) para la anterior m.a.s. de la v.a.  $\tilde{X} \otimes \tilde{X}$ .

En estas condiciones, asociada a la m. a. s.  $\{X^{(i)} : i = 1, \dots, N\}$ , y para un cierto  $n \in \mathbb{N}$  con  $n \leq N$ , se definen las siguientes vv.aa.

**Definición 1.6.1** Se define la media muestral, que denotaremos con  $\bar{X}_n$ , como la v.a. hilbertiana de  $L_2(\Omega^N, \mathcal{A}^N, P^N; H)$ , definida por

$$\bar{X}_n \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^n X^{(i)}$$

**Definición 1.6.2** Se define el operador de covarianza muestral, que denotaremos con  $\bar{C}_X^{(n)}$ , como la v.a. de  $L_1(\Omega^N, \mathcal{A}^N, P^N; HS)$ , definida por

$$\bar{C}_X^{(n)} \stackrel{\text{def}}{=} \frac{1}{n-1} \sum_{i=1}^n \widehat{X}^{(i)} \otimes \widehat{X}^{(i)}$$

siendo  $\widehat{X}^{(i)} \stackrel{\text{def}}{=} X^{(i)} - \bar{X}_n$ .

Por su parte, denotaremos con

$$\bar{K}_X^{(n)} \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^n \tilde{X}^{(i)} \otimes \tilde{X}^{(i)}$$

Obsérvese que la v.a.  $\bar{K}_X^{(n)}$  puede considerarse como la media muestral asociada a la m.a.s.,  $\{\tilde{X}^{(i)} \otimes \tilde{X}^{(i)} : i = 1, \dots, N\}$ , de la v.a.  $\tilde{X} \otimes \tilde{X}$ . Por esta razón, todo lo que se afirme para la v.a.  $\bar{X}_n$ , asociada a una m.a.s de la v.a.  $X$ , podrá ser extendible, bajo la imposición de convenientes condiciones, al caso de la v.a.  $\bar{K}_X^{(n)}$  para la correspondiente m.a.s. de la v.a.  $\tilde{X} \otimes \tilde{X}$ .

**Proposición 1.6.2** Supóngase que el espacio  $H$  es un espacio funcional, es decir, existe un conjunto de parámetros  $T$  tal que  $H \subseteq \mathbb{R}^T$ . Entonces, el operador de covarianza muestral  $\bar{C}_X^{(n)}$ , puede expresarse  $\forall f \in H$

$$C_X(f)(t) = \langle \hat{R}(t, *), f \rangle_H \quad \forall t \in T$$

siendo  $\hat{R} : T \times T \rightarrow \mathbb{R}$  la función de covarianza muestral, dada por

$$\hat{R}_w(t, s) \stackrel{\text{def}}{=} \frac{1}{n-1} \sum_{i=1}^n \widehat{X}_t^{(i)}(w) \widehat{X}_s^{(i)}(w) \in \mathbb{R}$$

denotando con  $\widehat{X}_t^{(i)}(w) \stackrel{\text{def}}{=} \widehat{X}^{(i)}(w)(t) = X^{(i)}(w)(t) - \bar{X}_n(w)(t)$

*Demostración*

Dado  $f \in H$ , utilizando la definición del operador de covarianza muestral se obtendría que

$$C_X(f) = \frac{1}{n-1} \sum_{i=1}^n \widehat{X}^{(i)} \otimes \widehat{X}^{(i)}(f) = \frac{1}{n-1} \sum_{i=1}^n \langle \widehat{X}^{(i)}, f \rangle_H \widehat{X}^{(i)}$$

De esta forma, para cada  $t \in T$  se obtendría que

$$C_X(f)(t) = \frac{1}{n-1} \sum_{i=1}^n \langle \widehat{X}^{(i)}, f \rangle_H \widehat{X}_t^{(i)} =$$

$$\frac{1}{n-1} \sum_{i=1}^n \langle \widehat{X}_t^{(i)} \widehat{X}^{(i)}, f \rangle_H = \langle \hat{R}(t, *), f \rangle_H$$

siendo  $\hat{R}(t, *) = \frac{1}{n-1} \sum_{i=1}^n \widehat{X}_t^{(i)} \widehat{X}^{(i)} \in H$ .

□

**Lema 1.6.1** *En las condiciones anteriores, se cumple:*

$$\mathbf{E}[\bar{X}_n] = \mu \quad y \quad \mathbf{E}[\bar{K}_X^{(n)}] = C_X$$

*Demostración*

Dado  $f \in H$ , arbitrario pero fijo, se tiene

$$\int_{\Omega^N} \langle \bar{X}_n(w), f \rangle_H dP^N(w) = \frac{1}{n} \sum_{i=1}^n \int_{\Omega^N} \langle X^{(i)}(w), f \rangle_H dP^N(w)$$

$$= \frac{1}{n} \sum_{i=1}^n \langle \mathbf{E}[X^{(i)}], f \rangle_H = \langle \mu, f \rangle_H$$

La demostración de la insesgadez del operador  $\bar{K}_X^{(n)}$  sigue el mismo esquema que la del Lema anterior, teniendo en cuenta que ahora se cumple:

$$\mathbf{E}[\tilde{X}^{(i)} \otimes \tilde{X}^{(i)}] = C_X \quad \forall i \in \{1, \dots, N\}$$

□

**Proposición 1.6.3** Si la v.a.  $X \in L_{2p}(\Omega, \mathcal{A}, P; H)$  (con  $p \in \mathbb{N}$ ), se cumpliría lo siguiente:

1. Si  $i \neq j$  se cumple  $X^{(i)} \otimes X^{(j)} \in L_{2p}(\Omega^N, \mathcal{A}^N, P^N; HS)$
2.  $X^{(i)} \otimes X^{(i)} \in L_p(\Omega^N, \mathcal{A}^N, P^N; HS)$
3. Si  $i \neq j$  se cumple  $\mathbf{E}[X^{(i)} \otimes X^{(j)}] = \mathbf{E}[X^{(i)}] \otimes \mathbf{E}[X^{(j)}]$
4.  $\mathbf{E}[X^{(i)} \otimes X^{(i)}] = C_x$

*Demostración*

En principio, dados  $i, j \in \{1, \dots, N\}$  y  $w = (w_1, \dots, w_N) \in \Omega^N$ , se cumple:

$$\|X^{(i)} \otimes X^{(j)}(w)\|_{HS} = \|X(w_i) \otimes X(w_j)\|_{HS} = \|X(w_i)\|_H \|X(w_j)\|_H$$

Teniendo en cuenta esto se obtiene

$$\mathbf{E}[\|X^{(i)} \otimes X^{(j)}\|_{HS}^v] = \int_{\Omega^N} \|X(w_i)\|_H^v \|X(w_j)\|_H^v dP^N(w)$$

con lo que, aplicando el Teorema de Fubini, pueden distinguirse las dos siguientes situaciones:

- $\mathbf{E}[\|X^{(i)} \otimes X^{(i)}\|_{HS}^p] = \mathbf{E}[\|X\|_H^{2p}] < \infty$
- $\mathbf{E}[\|X^{(i)} \otimes X^{(j)}\|_{HS}^{2p}] = (\mathbf{E}[\|X\|_H^{2p}])^2 < \infty$  Si  $i \neq j$

Con objeto de probar el tercer resultado ( $i \neq j$ ), consideraremos una base ortonormal,  $\mathcal{B} = \{e_k : k \in K\}$ , del espacio de Hilbert  $(H, \langle \cdot, \cdot \rangle_H)$ . En concreto, para el operador  $e_k \otimes e_l \in HS$  se cumple:

$$\begin{aligned} \langle \mathbf{E}[X^{(i)} \otimes X^{(j)}], e_k \otimes e_l \rangle_{HS} &= \int_{\Omega^N} \langle X^{(i)} \otimes X^{(j)}(w), e_k \otimes e_l \rangle_{HS} dP^N(w) \\ &= \int_{\Omega^N} \langle X(w_i) \otimes X(w_j)(e_k), e_l \rangle_H dP^N(w) = \\ &= \int_{\Omega^2} \langle X(w_i), e_k \rangle_H \langle X(w_j), e_l \rangle_H dP(w_i) dP(w_j) = \end{aligned}$$

$$\begin{aligned}
&= \langle \mathbf{E}[X], e_k \rangle_H \langle \mathbf{E}[X], e_l \rangle_H = \langle \langle \mathbf{E}[X], e_k \rangle_H \mathbf{E}[X], e_l \rangle_H \\
&= \langle \mathbf{E}[X^{(i)}] \otimes \mathbf{E}[X^{(j)}](e_k), e_l \rangle_H = \langle \mathbf{E}[X^{(i)}] \otimes \mathbf{E}[X^{(j)}], e_k \otimes e_l \rangle_{HS}
\end{aligned}$$

De lo anterior, y teniendo en cuenta que el subespacio vectorial de operadores de rango finito,  $\text{Lin}\{e_k \otimes e_l : k, l \in K\}$ , es denso en  $HS$  puede afirmarse que

$$\langle \mathbf{E}[X^{(i)} \otimes X^{(j)}], G \rangle = \langle \mathbf{E}[X^{(i)}] \otimes \mathbf{E}[X^{(j)}], G \rangle_{HS} \quad \forall G \in HS$$

□

**Proposición 1.6.4** Siendo  $\widetilde{X}_n \stackrel{\text{def}}{=} \bar{X}_n - \mu$ , se cumple que

$$C_{\widetilde{X}_n} \stackrel{\text{def}}{=} \mathbf{E}[\widetilde{X}_n \otimes \widetilde{X}_n] = \frac{1}{n} C_X$$

*Demostración*

En principio, obsérvese que por ser  $\widetilde{X}_n = \frac{1}{n} \sum_{i=1}^n \tilde{X}^{(i)}$  se obtiene que

$$\widetilde{X}_n \otimes \widetilde{X}_n = \frac{1}{n^2} \sum_{i \neq j} \tilde{X}^{(i)} \otimes \tilde{X}^{(j)} + \frac{1}{n^2} \sum_{i=1}^n \tilde{X}^{(i)} \otimes \tilde{X}^{(i)}$$

de donde aplicando esperanzas, y teniendo en cuenta que

$$\mathbf{E}[\tilde{X}^{(i)} \otimes \tilde{X}^{(j)}] = \mathbf{E}[\tilde{X}^{(i)}] \otimes \mathbf{E}[\tilde{X}^{(j)}] = 0$$

se obtiene lo siguiente:

$$\mathbf{E}[\widetilde{X}_n \otimes \widetilde{X}_n] = \frac{1}{n^2} \sum_{i=1}^n \mathbf{E}[\tilde{X}^{(i)} \otimes \tilde{X}^{(i)}] = \frac{1}{n} C_X$$

□

**Lema 1.6.2**

$$(n-1) \bar{C}_X^{(n)} = n \bar{K}_X^{(n)} - n \widetilde{X}_n \otimes \widetilde{X}_n$$

*Demostración*

En primer lugar, obsérvese que, dado  $i \in \{1, \dots, N\}$ , se cumple

$$\begin{aligned} \widehat{X}^{(i)} \otimes \widehat{X}^{(i)} &= [(X^{(i)} - \mu) - (\bar{X}_n - \mu)] \otimes [(X^{(i)} - \mu) - (\bar{X}_n - \mu)] = \\ &= \tilde{X}^{(i)} \otimes \tilde{X}^{(i)} - \tilde{X}^{(i)} \otimes \widetilde{\bar{X}}_n - \widetilde{\bar{X}}_n \otimes \tilde{X}^{(i)} + \widetilde{\bar{X}}_n \otimes \widetilde{\bar{X}}_n = \end{aligned}$$

Teniendo en cuenta lo anterior, se obtiene

$$\begin{aligned} \sum_{i=1}^n \widehat{X}^{(i)} \otimes \widehat{X}^{(i)} &= \sum_{i=1}^n \tilde{X}^{(i)} \otimes \tilde{X}^{(i)} - n \widetilde{\bar{X}}_n \otimes \widetilde{\bar{X}}_n - n \widetilde{\bar{X}}_n \otimes \widetilde{\bar{X}}_n + \\ &+ n \widetilde{\bar{X}}_n \otimes \widetilde{\bar{X}}_n = n \bar{K}_X^{(n)} - n \widetilde{\bar{X}}_n \otimes \widetilde{\bar{X}}_n \end{aligned}$$

□

**Corolario 1.6.1** *En las condiciones anteriores, se cumple:*

1.  $\bar{C}_X^{(n)} \in L_1(\Omega^N, \mathcal{A}^N, P^N; HS)$
2.  $\mathbf{E}[\bar{C}_X^{(n)}] = C_X$

*Demostración*

Es consecuencia, de aplicar lo mostrado hasta el momento y del Lema anterior.

$$(n-1) \mathbf{E}[\bar{C}_X^{(n)}] = n \mathbf{E}[\bar{K}_X^{(n)}] - n \mathbf{E}[\widetilde{\bar{X}}_n \otimes \widetilde{\bar{X}}_n] = n C_X - C_X$$

□

### 1.6.2 ACP ESTIMADO

Uno de los primeros trabajos, donde se estudia el problema de la estimación de los elementos propios del operador de covarianza, es debido a Deville (1973b). En dicho trabajo además de proponer estimadores para dichos elementos propios, se obtienen sus *momentos*, y además se analiza su comportamiento asintótico.

La definición de dichos estimadores viene dada a partir de los elementos propios muestrales de  $\bar{C}_X^{(n)}$ , los cuales vendrán dados, para cada  $w \in \Omega^N$ , como soluciones de la ecuación siguiente:

$$\bar{C}_X^{(n)}(w) \left( \hat{f}_i^{(n)}(w) \right) = \hat{\lambda}_i^{(n)}(w) \hat{f}_i^{(n)}(w)$$

bajo la restricción de que  $\{\hat{f}_i^{(n)}(w) : i = 1, \dots, N\}$  sea un sistema ortonormal del espacio  $H$ . Consideraremos que los valores propios muestrales han sido ordenados, según el subíndice, como sigue:

$$\hat{\lambda}_i^{(n)} \geq \hat{\lambda}_j^{(n)} \quad \text{Si } i < j$$

Una forma equivalente de interpretar el proceso de estimación asociado es la siguiente. Considerados los  $n$  elementos obtenidos a partir de una m.a.s. de tamaño  $n$  de la v.a. hilbertiana  $X$

$$\{X(w_i^*) : i = 1, \dots, n\} \subseteq H,$$

la obtención del ACP estimado de  $X$  sería equivalente a considerar el ACP de  $x$  (restricción de  $X$ ) sobre el espacio probabilístico  $(\Omega^*, \mathcal{P}(\Omega^*), P^*)$ , siendo  $\Omega^* \stackrel{\text{def}}{=} \{w_1^*, \dots, w_n^*\}$ , y para cada  $\forall l = 1, \dots, n$

$$x(w_l^*) \stackrel{\text{def}}{=} X(w_l^*) \quad P^*(\{w_l^*\}) \stackrel{\text{def}}{=} n^{-1}$$

Puede comprobarse, notando con  $w^* = (w_1^*, \dots, w_n^*)$ , que se obtiene

$$\mathbf{E}[x] = \bar{X}_n(w^*) \quad \mathcal{C}_x = \bar{C}_X^{(n)}(w^*)$$

Tras obtener los elementos propios del operador de covarianza  $\mathcal{C}_x$

$$\mathcal{C}_x \left( \hat{f}_i^* \right) = \hat{\lambda}_i^* \hat{f}_i^*,$$

y ordenados en orden decreciente con respecto al orden del índice, se obtendría la siguiente descomposición del operador  $\mathcal{C}_x$

$$\mathcal{C}_x = \sum_{i=1}^n \hat{\lambda}_i^* \hat{f}_i^* \otimes \hat{f}_i^*$$

siendo  $\hat{f}_i^{(n)}(w^*) \stackrel{\text{def}}{=} \hat{f}_i^*$  y  $\hat{\lambda}_i^{(n)}(w^*) \stackrel{\text{def}}{=} \hat{\lambda}_i^*$ .



Con la anterior descomposición, y teniendo en cuenta la correspondiente descomposición del operador de covarianza,  $C_X$ , de la v.a. hilbertiana  $X$ , dada por

$$C_X = \sum_{i \in I} \lambda_i f_i \otimes f_i$$

con los valores propios ordenados en forma decreciente

$$\lambda_i \geq \lambda_j \quad \text{si } i < j$$

y apareciendo tantas veces como indique su multiplicidad, se establecen las definiciones de los estimadores siguientes.

**Definición 1.6.3** (Valores principales muestrales)

- En el caso de que todos los valores propios no nulos de  $C_X$  sean de multiplicidad 1 (simple), se define el estimador del  $i$ -ésimo mayor valor propio de  $C_X$ ,  $\lambda_i$ , a través del  $i$ -ésimo mayor valor propio muestral de  $\bar{C}_X^{(n)}$ ,  $\hat{\lambda}_i^{(n)}$ .
- En el caso de que alguno de los valores propios no nulos de  $C_X$ ,  $\lambda$ , tenga multiplicidad mayor que 1, la estimación de dicho valor propio que se suele utilizar es el valor medio de las estimaciones proporcionados por los valores propios muestrales de  $\bar{C}_X^{(n)}$  siguientes

$$\{\hat{\lambda}_i^{(n)} : i \text{ tal que } \lambda_i = \lambda\}$$

Por su parte, en relación a la estimación de los factores principales ha de tenerse presente que *no están determinados de forma única*. Por ejemplo, aún en el caso de que  $\lambda_i$  sea simple,  $\hat{f}_i^{(n)}$  podría ser el estimador de  $f_i$ , o en su caso  $-f_i$ .

**Definición 1.6.4** Denominaremos a  $\hat{f}_i^{(n)}$  como el  $i$ -ésimo factor principal muestral.

Las mismas consideraciones, hechas para los factores principales muestrales, caben hacerlas para las correspondientes componentes principales muestrales, las cuales vendrían dadas por

$$\hat{\psi}_i^{(n)}(w_i^*) \stackrel{\text{def}}{=} \langle x(w_i^*), f_i^* \rangle_H \quad \forall i = 1, \dots, n$$

Una vez obtenidos los anteriores elementos muestrales se obtendría

$$x(w_l) = X(w_l) = \sum_{i=1}^n \hat{\psi}_i^{(n)}(w_l) \hat{f}_i^* \quad \forall l = 1, \dots, n$$

Obsérvese que para una m.a.s. de tamaño  $n$ , teniendo en cuenta que las estimaciones de  $\bar{C}_X^{(n)}$  son operadores de rango finito, son nulos, para cada  $i > n$ . las estimaciones correspondientes de  $\lambda_i$ ,  $f_i$ , y  $\psi_i$ , es decir,

$$\hat{\psi}_i^{(n)} = 0 \quad \hat{\lambda}_i^{(n)} = 0 \quad \hat{f}_i^{(n)}$$

En este ambiente, un estimador natural de la variabilidad total,  $V \stackrel{\text{def}}{=} \text{tr}(\mathcal{C}_X)$ , vendría dado por

$$\hat{V} \stackrel{\text{def}}{=} \text{tr}(\bar{C}_X^{(n)}) = \sum_{i=1}^n \hat{\lambda}_i^{(n)}$$

**Proposición 1.6.5**  $\hat{V}$  es un estimador insesgado de  $V$ .

*Demostración*

Dada una base ortonormal  $\mathcal{B} = \{e_k : k \in K\}$  de  $H$ , sabemos que

$$\hat{V} = \sum_{k \in K} \langle \bar{C}_X^{(n)}(e_k), e_k \rangle_H = \sum_{k \in K} \langle \bar{C}_X^{(n)}, e_k \otimes e_k \rangle_{HS}$$

de donde

$$\mathbf{E}[\hat{V}] = \sum_{k \in K} \mathbf{E}[\langle \bar{C}_X^{(n)}(e_k), e_k \rangle_H] = \sum_{k \in K} \mathbf{E}[\langle \bar{C}_X^{(n)}, e_k \otimes e_k \rangle_{HS}]$$

Con lo que haciendo uso de la insesgadez del estimador  $\bar{C}_X^{(n)}$  se obtiene

$$\mathbf{E}[\hat{V}] = \sum_{k \in K} \langle \mathcal{C}_X, e_k \otimes e_k \rangle_{HS} = V$$

□

### 1.6.3 ALGUNOS RESULTADOS ASINTOTICOS

Continuando con la notación anterior, supondremos que la m.a.s. está formada por una colección infinita numerable de v.a. hilbertianas ( $N = \infty$ )

$$\{X^{(i)} : i \in \mathbb{N}\}$$

definidas sobre el espacio probabilístico  $(\Omega^\infty, \mathcal{A}^\infty, P^\infty)$ .

Uno de los primeros resultados que puede obtenerse es el siguiente:

**Proposición 1.6.6** *En las condiciones anteriores, la sucesión de medias muestrales,  $\{\bar{X}_n/n \in \mathbb{N}\} \subseteq L_2(\Omega^\infty, \mathcal{A}^\infty, P^\infty; H)$ , verifica:*

1.  $\lim_{n \rightarrow \infty} \bar{X}_n = \mu$  ( en el espacio  $L_2(\Omega^\infty, \mathcal{A}^\infty, P^\infty; H)$  )
2. c.t.  $w \in \Omega^\infty [P^\infty]$   $\lim_{n \rightarrow \infty} \bar{X}_n(w) = \mu$  ( en el espacio  $(H, \langle *, * \rangle_H)$  )
3.  $\bar{X}_n$  es un estimador consistente de  $\mu$ .

*Demostración*

Una demostración de este resultado puede verse en Deville (1973b).

□

Para los resultados asintóticos, que vamos a enunciar a continuación, es necesario exigir como condición adicional que la v.a.  $X$  cumple lo siguiente:

$$X \in L_4(\Omega, \mathcal{A}, P; H)$$

**Proposición 1.6.7** *La sucesión de covarianzas muestrales,*

$$\{\bar{C}_X^{(n)}/n \in \mathbb{N}\} \subseteq L_2(\Omega^\infty, \mathcal{A}^\infty, P^\infty; HS), \text{ cumple:}$$

1.  $\lim_{n \rightarrow \infty} \bar{C}_X^{(n)} = C_X$  ( en el espacio  $L_2(\Omega^\infty, \mathcal{A}^\infty, P^\infty; HS)$  )
2. c.t.  $w \in \Omega^\infty [P^\infty]$   $\lim_{n \rightarrow \infty} \bar{C}_X^{(n)}(w) = C_X$  ( en el espacio  $HS$  )
3.  $\bar{C}_X^{(n)}$  es un estimador consistente de  $C_X$ .

*Demostración*

Una demostración de este resultado puede verse en Deville (1973b). □

**Proposición 1.6.8** *La sucesión,*

$$\{\hat{V}^{(n)} \stackrel{\text{def}}{=} \text{tr}(\bar{C}_X^{(n)})/n \in \mathbb{N}\} \subseteq L_2(\Omega^\infty, \mathcal{A}^\infty, P^\infty; \mathbb{R}),$$

*converge en m.c., probabilidad y c.s a V.*

*Demostración*

Una demostración de este resultado puede verse en Deville (1973b). □

**Teorema 1.6.1** *Dado  $i \in I$ , se cumple que*

$$\{\hat{\lambda}_i^{(n)}\}_{n \in \mathbb{N}} \subseteq L_2(\Omega^\infty, \mathcal{A}^\infty, P^\infty; \mathbb{R}).$$

*Además,*

- *si todos los valores propios no nulos de  $C_X$  son simples, entonces:*

$$\lim_{n \rightarrow \infty} \hat{\lambda}_i^{(n)} = \lambda_i \quad \text{en m.c., probabilidad y c.s.}$$

- *si  $\lambda_i$  un valor propio no nulo de  $C_X$ , y considerando*

$$I_i = \{j \in I : \lambda_j = \lambda_i\} \quad \text{card}(I_i) = k_i$$

*entonces, las  $k_i$  sucesiones para cada  $j \in I_i$*

$$\lim_{n \rightarrow \infty} \hat{\lambda}_j^{(n)} = \lambda_i \quad \text{probabilidad y c.s.}$$

*Demostración* La demostración del primer resultado puede verse en Deville (1973b), y la del segundo resultado en Dauxois (1982). □

**Teorema 1.6.2** *Dado  $i \in I$ , se cumple que  $\{\hat{f}_i^{(n)}\}_{n \in \mathbb{N}} \subseteq L_2(\Omega^\infty, \mathcal{A}^\infty, P^\infty; H)$ .*

*Además,*

- si todos los valores propios no nulos de  $C_X$  son simples, en el caso de que exista  $n_o \in \mathbb{N}$  tal que si  $n \geq n_o$

$$\langle \hat{f}_i^{(n)}, f_i \rangle_H \geq 0 \quad (\text{c.s.})$$

se cumple  $\lim_{n \rightarrow \infty} \hat{f}_i^{(n)} = f_i$  c.s., m.c. y probabilidad

- las proyecciones muestrales

$$P_i^{(n)} = \sum_{j \in I_i} \hat{f}_j^{(n)} \otimes \hat{f}_j^{(n)}$$

convergen, cuando  $n \rightarrow \infty$ , en probabilidad y c.s., a  $P_i$ . Siendo,  $P_i^{(n)}(w)$  la proyección ortogonal sobre el subespacio engendrado por

$$\{\hat{f}_j^{(n)}(w) : j \in I_i\},$$

y  $P_i$  la proyección ortogonal sobre  $\text{Ker}(C_X - \lambda_i I_H)$ .

#### Demostración

El primer resultado puede verse en Deville (1973b), y el segundo en Dauxois (1982).

□



## Capítulo 2

# GEOMETRIAS ALTERNATIVAS EN EL ACP DE UNA V.A. HILBERTIANA

### §2.1 ACP FILTRANTE

El objeto de esta sección es mostrar un caso en donde se plantea el ACP de una v.a. hilbertiana con respecto a una determinada geometría. En concreto, se analizará el caso de la realización de un ACP de datos filtrados, a través de un cierto filtro lineal; mostrándose como caso particular de ACP de los datos originales con respecto a una geometría determinada, entre otros, por el filtro.

Consideremos un espacio de Hilbert  $(G, \tau)$ , y un espacio vectorial,  $H$ , sobre el que tenemos definido un operador lineal e inyectivo

$$F : H \mapsto (G, \tau)$$

Definimos la aplicación,  $\rho : H \times H \mapsto \mathbb{R}$ , como sigue:

$$\rho(h_1, h_2) \stackrel{\text{def}}{=} \tau(F(h_1), F(h_2)) \quad \forall h_1, h_2 \in H$$

De esta forma, dotamos al espacio vectorial  $H$  de una estructura geométrica,  $(H, \rho)$ , inducida por el espacio de Hilbert  $(G, \tau)$  a través del operador  $F$ .

Obsérvese que, teniendo en cuenta la notación introducida, el espacio  $H$  contendrá los *datos originales* y el filtro lineal vendrá determinado por el operador  $F$ .

### 2.1.1 ANALISIS DE LA ESTRUCTURA GEOMETRICA

**Proposición 2.1.1** *En las condiciones anteriores, y si  $F(H)$  tiene estructura de espacio vectorial son equivalentes las siguientes afirmaciones:*

1.  $\rho$  es un producto escalar sobre  $H$ .
2.  $F$  es un operador lineal e inyectivo.

□

**Proposición 2.1.2** *El operador  $F$  definido*

$$F : (H, \rho) \longrightarrow (F(H), \tau) \subseteq (G, \tau)$$

*es una isometría lineal.*

*Demostración*

Obsérvese que se cumple:

$$\tau(F(h_1), F(h_2)) = \rho(h_1, h_2) \quad \forall h_1, h_2 \in H$$

□

Es decir, los espacios prehilbertianos  $(H, \rho)$  y  $(F(H), \tau)$  son *isométricamente iguales*. Y en definitiva, la estructura geométrica del espacio  $(H, \rho)$  puede considerarse isométricamente contenida en el espacio  $(G, \tau)$ .

**Corolario 2.1.1** *Si  $F(H) = G$  entonces  $(H, \rho)$  posee estructura de espacio de Hilbert.*

**Proposición 2.1.3** *En el caso de que exista,  $F^*$ , el operador adjunto del operador*

$$F : (H, \rho) \longrightarrow (F(H), \tau) \subseteq (G, \tau)$$

*se obtiene que  $F^* \circ F = I_H$ .*

*Demostración*

Dado  $g = F(f) \in F(H)$ , arbitrario pero fijo, se cumple:

$$\tau(g, F(h)) = \tau(F(f), F(h)) = \rho(f, h) \quad \forall h \in H$$

de donde se concluye que  $g \in \text{Dom}(F^*)$ , siendo además

$$F^*(g) = f = F^* \circ F(f)$$

Es decir, se obtiene que  $F(H) \subseteq \text{Dom}(F^*)$ , y además  $F^* \circ F = I_H$ .



□

Obsérvese que en definitiva el resultado anterior establece que

$$F^* = F^{-1} \quad \text{sobre el conjunto } F(H)$$

Una situación en la que quedaría garantizada la existencia del operador adjunto  $F^*$  sería, por ejemplo, en el caso en que el espacio  $(H, \rho)$  fuese un espacio de Hilbert. En este caso, al ser  $F$  un operador lineal y continuo, entre dos espacio de Hilbert, se tendría asegurada la existencia de dicho operador adjunto

$$F^* : (G, \tau) \longrightarrow (H, \rho) .$$

### 2.1.2 ESTUDIO DEL ACP FILTRANTE

Sea  $X : (\Omega, \mathcal{A}, P) \longrightarrow H$  una aplicación con valores en el espacio  $H$ .

Supondremos que  $(H, \rho)$  posee estructura de espacio de Hilbert. Esto originará, por lo establecido en la Proposición (2.1.2), que  $(F(H), \tau)$  también posea estructura hilbertiana. Así mismo, supondremos que  $X$  es  $\mathcal{B}_{(H, \rho)}/\mathcal{A}$ -medible. Con lo cual, según la Proposición (2.1.2), puede asegurarse que  $F \circ X$  resulta ser  $\mathcal{B}_{(G, \tau)}/\mathcal{A}$ -medible.

El problema que aquí nos planteamos es analizar la forma de obtener el ACP de  $X$  con respecto a la geometría  $(H, \rho)$ , a partir de la estructura geométrica de referencia  $(G, \tau)$ . Obsérvese que la idea que motiva esta pretensión es el hecho de que el espacio  $(H, \rho)$  puede considerarse *isométricamente incluido* en  $(G, \tau)$ .

En estas condiciones, consideraremos la v.a. hilbertiana

$$FX \stackrel{\text{def}}{=} F \circ X : (\Omega, \mathcal{A}, P) \longrightarrow (F(H), \tau) \subseteq (G, \tau)$$

Para simplificar la notación, siempre que no se especifique producto escalar se estará haciendo referencia a  $\tau$ .

**Proposición 2.1.4** *Dado  $p \in \mathbb{R} - \{0\}$  se cumple:*

*$X \in L_p(\Omega, \mathcal{A}, P; H, \rho)$ , si y solamente si,  $FX \in L_p(\Omega, \mathcal{A}, P; G, \tau)$ .*

*Demostración*

Obsérvese que, para cada  $w \in \Omega$ , se cumple:

$$\|FX(w)\|_{\tau}^p = \tau(F(X(w)), F(X(w)))^{\frac{p}{2}} = \|X(w)\|_{\rho}^p$$

□

**Proposición 2.1.5**

Si la v.a.  $X \in L_1(\Omega, \mathcal{A}, P; H, \rho)$ , se tiene que  $\mathbf{E}[FX] \in F(H)$  de forma que

$$F(\mathbf{E}^\rho[X]) = \mathbf{E}[FX]$$

*Demostración*

En primer lugar, por ser  $FX$  una v.a. con valores en el espacio de Hilbert  $(F(H), \tau)$ , puede garantizarse que  $\mathbf{E}[FX] \in F(H)$ .

Así mismo, haciendo uso de la definición de esperanza de una v.a. hilbertiana se razona, para cada  $h \in H$ , como sigue:

$$\begin{aligned} \tau(\mathbf{E}[FX], F(h)) &= \int_{\Omega} \tau(F(X(w)), F(h)) dP(w) = \int_{\Omega} \rho(X(w), h) dP(w) = \\ &= \rho(\mathbf{E}^\rho[X], h) = \tau(F(\mathbf{E}^\rho[X]), F(h)) \end{aligned}$$

de donde concluimos que  $F(\mathbf{E}^\rho[X]) = \mathbf{E}[FX]$ .

□

**Corolario 2.1.2**

Si denotamos con  $\widetilde{FX} = FX - \mathbf{E}[FX]$  y con  $\tilde{X} = X - \mathbf{E}^\rho[X]$  se cumple:

$$\widetilde{FX} = F(\tilde{X})$$

*Demostración*

Obsérvese que haciendo uso del resultado anterior, para cada  $w \in \Omega$ , se obtiene:

$$\widetilde{FX}(w) = FX(w) - \mathbf{E}[FX] = F(X(w)) - F(\mathbf{E}^\rho[X]) = F(\tilde{X}(w))$$

□

**Proposición 2.1.6** Si  $X \in L_2(\Omega, \mathcal{A}, P; H, \rho)$  se obtiene:

$$\begin{aligned} U_x^\rho &= U_{FX} \circ F & U_x^{\rho*} &= F^{-1} \circ U_{FX}^* \\ C_x^\rho &= F^{-1} \circ C_{FX} \circ F & W_x^\rho &= W_{FX} \end{aligned}$$

*Demostración*

Comenzaremos analizando el operador  $U_x^\rho : (H, \rho) \mapsto L_2(\Omega, \mathcal{A}, P)$ . Haciendo uso de su definición, dado  $h \in H$  arbitrario pero fijo, se razona como sigue:

$$\begin{aligned} U_x^\rho(h)(w) &= \rho(\tilde{X}(w), h) = \tau(F(\tilde{X}(w)), F(h)) = \\ &= \tau(\widetilde{FX}(w), F(h)) = U_{FX}(F(h))(w) \quad \forall w \in \Omega \end{aligned}$$

de donde se concluye que  $U_x^\rho(h) = U_{FX} \circ F(h) \quad \forall h \in H$

Obsérvese que los demás resultados son consecuencia del demostrado y de aplicar las definiciones de los correspondientes operadores sin más que tener en cuenta la Proposición (2.1.3), la cuál establece que  $F^*$ , sobre el subespacio  $F(H)$ , coincide con  $F^{-1}$ .

□

**Lema 2.1.1** *Los espectros de los operadores  $C_x^\rho$  y  $C_{FX}$  son iguales. Además, la relación entre sus subespacios propios es la siguiente:*

$$F(E_\lambda) = V_\lambda$$

donde para cada  $\lambda$  perteneciente al espectro puntual

$$E_\lambda \stackrel{\text{def}}{=} \text{Ker}(C_x^\rho - \lambda I_H) \quad V_\lambda \stackrel{\text{def}}{=} \text{Ker}(C_{FX} - \lambda I_{F(H)})$$

*Demostración*

Dados  $\lambda \in P\sigma(C_x^\rho)$  y  $h \in E_\lambda - \{0\}$  se cumple

$$C_x^\rho(h) = \lambda h = F^{-1} \circ C_{FX} \circ F(h)$$

Aplicando el operador  $F$  a la expresión anterior se obtiene

$$C_{FX}(F(h)) = \lambda F(h)$$

de donde podemos concluir que  $\lambda \in P\sigma(C_{FX})$ , ya que  $F(h) \neq 0$ , con lo que  $F(h) \in V_\lambda$ .

De lo anterior podemos afirmar que en principio se tiene:

$$P\sigma(C_x^\rho) \subseteq P\sigma(C_{FX}) \quad \text{y} \quad F(E_\lambda) \subseteq V_\lambda$$

Recíprocamente, dados  $\lambda \in P\sigma(C_{FX})$  y  $g \in V_\lambda - \{0\}$ , obsérvese que al ser  $g \in F(H)$ , existirá un único  $h \in H - \{0\}$  cumpliendo  $g = F(h)$ . Con ello, se obtiene

$$C_{FX}(g) = \lambda g \implies C_{FX} \circ F(h) = F(\lambda h)$$

que por la inyectividad de  $F$  se concluye

$$F^{-1} \circ C_{FX} \circ F(h) = C_x^\rho(h) = \lambda h$$

de donde se obtiene que  $\lambda \in P\sigma(C_x^\rho)$  y  $h \in V_\lambda$ .

De esta forma, podemos concluir

$$P\sigma(C_{FX}) \subseteq P\sigma(C_x^\rho) \quad \text{y} \quad V_\lambda \subseteq F(E_\lambda)$$

□

**Teorema 2.1.1** *Teniendo en cuenta la notación introducida en el Lema anterior, el ACP de la v.a.  $X$  respecto de la geometría  $(H, \rho)$  es equivalente al ACP de la v.a.  $FX$  respecto a la geometría  $(G, \tau)$ . Es decir, el ACP de uno de ellos determina un ACP del otro, de la forma siguiente:*

$$X(w) = \sum_{i \in I} f_i \psi_i(w)$$

$$FX(w) = \sum_{i \in I} F(f_i) \psi_i(w)$$

cumpléndose para cada  $\lambda \in P\sigma(C_x^\rho)$  lo siguiente:

$$V_\lambda = F(E_\lambda)$$

*Demostración*

En principio, en el Lema anterior se establece la relación existente entre los espectros y entre los subespacios propios asociados a los correspondientes operadores de covarianza.

Con objeto de mostrar que cualquier representación en componentes principales de  $X$  tiene asociado de forma única una representación en componentes principales, y recíprocamente, obsérvese que se tienen los siguientes resultados:

$$\tau(F(f_i), F(f_j)) = \rho(f_i, f_j) \quad \forall i, j \in I$$

$$U_x^\rho(f_i) = U_{FX} \circ F(f_i) = U_{FX}(F(f_i)) \quad \forall i \in I$$

□

Una primera consecuencia que podemos extraer de este resultado es que la estructura aleatoria, dada por las correspondientes componentes principales, es idéntica en ambos ACP. La única diferencia se establece en relación a la estructura geométrica a través de  $F$ .

**Proposición 2.1.7** *En las condiciones del Teorema anterior, se cumple:*

$$\text{tr}(C_x^\rho) = \text{tr}(C_{FX})$$

*Demostración*

Obsérvese que utilizando lo obtenido en la Proposición (2.1.6) se obtiene:

$$\text{tr}(C_x^\rho) = \text{tr}(W_x^\rho) = \text{tr}(W_{FX}) = \text{tr}(C_{FX})$$

□

### 2.1.3 ACP DEL INPUT DE UN SISTEMA ESTOCÁSTICO LINEAL

#### INTRODUCCION

En algunos trabajos, en especial en el campo de la psicología, se utiliza la técnica del ACP con objeto de *interpretar* el efecto del input externo de un sistema sobre el correspondiente output; en especial para casos en los que el output recoja la evolución de una cierta medida a lo largo de un intervalo de tiempo,  $T \subseteq \mathbb{R}$  (datos funcionales). En esta sección mostraremos cómo lo allí realizado puede interpretarse en términos de un ACP *filtrante* sobre el espacio de outputs.

Intuitivamente, la idea es realizar una *descomposición* en el espacio funcional que contiene al conjunto de posibles outputs del sistema, de forma que se obtenga: por un lado, una parte que de alguna forma pueda ser debida a la estructura interna del sistema, y por otro lado, la parte que pueda ser justificada *sólo* por la acción del input externo en cuestión. Como más adelante se mostrará, la forma de llevar a cabo esta descomposición es dotar al espacio funcional output con una estructura geométrica, de forma que en dicha estructura aparezcan como dos subespacios funcionales ortogonales los debidos al sistema y al input externo.

Concretamente, en trabajos como los de Besse et al. (1986) y Ramsay et al. (1995), a partir de un cierto conjunto de outputs observados del sistema (funciones)

$$\{x_i : T \mapsto \mathbb{R} : i = 1, \dots, n\}$$

se establece como espacio de outputs del sistema al espacio funcional (para un cierto  $m \in \mathbb{N}$ ):

$$W^m = \left\{ f : T \longrightarrow \mathbb{R} : \begin{array}{l} f, f', \dots, f^{(m-1)} \text{ absolutamente continuas} \\ f^{(m)} \in L_2(T) \end{array} \right\}$$

Así mismo, mediante argumentos de tipo psicológico, se identifica la estructura del sistema a través de la determinación de un cierto operador diferencial  $L : W^m \longrightarrow L_2(T)$ , de la forma siguiente:

$$L = \sum_{i=0}^m a_i D^{(i)} \quad (\text{con } a_i : T \longrightarrow \mathbb{R} \text{ continua})$$

En esencia, la determinación del operador  $L$  se realiza con la idea de que  $Ker(L)$  describa cualquier comportamiento posible del sistema bajo la ausencia de influencia externa.

Así mismo, se considera una familia de funcionales lineales

$$B_j : W^m \rightarrow \mathbb{R} \quad \forall j \in \{1, \dots, m\}$$

a los cuales se les exige que, para cualesquiera  $\alpha = (\alpha_1, \dots, \alpha_m)^t \in \mathbb{R}^m$  y  $u \in L_2(T)$ , el problema de valores iniciales (PVI), dado por

$$\begin{cases} L(x) = u \\ B_j(x) = \alpha_j \quad \forall j = 1, \dots, m \end{cases} \quad (\text{condiciones iniciales})$$

posea solución única.

De esta forma, cada output  $x_i$  constituiría la única solución del problema

$$\begin{cases} L(x) = u_i & \text{con } u_i \stackrel{\text{def}}{=} L(x_i) \text{ (input externo)} \\ B(x) = B(x_i) \in \mathbb{R}^m & B(x) \stackrel{\text{def}}{=} (B_1(x), \dots, B_m(x))^t \end{cases}$$

con lo que podría descomponerse en la forma siguiente:

$$x_i = x_i^{(1)} + x_i^{(2)}$$

<i>estructura interna del sistema</i>	<i>efecto del input externo</i>
$L(x_i^{(1)}) = 0$	$L(x_i^{(2)}) = u_i$
$B(x_i^{(1)}) = B(x_i)$	$B(x_i^{(2)}) = 0$

pudiendo interpretarse los sumandos de la forma siguiente:

- $x_i^{(1)}$  describe la evolución debida sólo al sistema, es decir, constituiría la evolución observada del sistema en el caso de que no hubiese actuado el input  $u_i$  partiendo de idénticas condiciones iniciales.
- $x_i^{(2)}$  describe la evolución del sistema debida al input.

Con objeto de interpretar los efectos del input externo, en los trabajos citados anteriormente se realiza el ACP clásico ( en  $L_2(T)$  ) del conjunto de funciones inputs

$$\{u_i = L(x_i) : i = 1, \dots, m\}.$$

A continuación, intentaremos mostrar que esta metodología puede considerarse, de alguna forma, como un caso particular de ACP en el espacio funcional de outputs, respecto de una geometría definida a tal efecto para el problema en cuestión. Además, como a continuación se analizará, este esquema podría generalizarse a otras situaciones en las que el operador no fuese diferencial.

### ESTABLECIMIENTO DE UNA GEOMETRIA

Consideremos un espacio vectorial  $W$ , en general de dimensión infinita, y sobre el que se encuentran definidos:

- un operador lineal,  $L : W \rightarrow S$ , con  $\dim Ker(L) = m$ , siendo  $(S, \langle *, * \rangle_S)$  un espacio de Hilbert.
- una familia de  $m$ -funcionales lineales  $\{B_j : W \rightarrow \mathbb{R} : j = 1, \dots, m\}$

verificando las siguientes condiciones:

1.  $L$  es sobreyectiva.
2.  $Ker(L) \cap Ker(B) = \{0\}$

siendo  $B : W \rightarrow \mathbb{R}^m$  la aplicación lineal dada por

$$B(x) \stackrel{\text{def}}{=} (B_1(x), \dots, B_m(x))^t \quad \forall x \in W$$

con lo que  $Ker(B) = \bigcap_{i=1}^m Ker(B_i)$ .

En esta situación, consideremos los subespacios vectoriales de  $W$ , a los que denotaremos con:

$$W_1 \stackrel{\text{def}}{=} Ker(L) \quad W_2 \stackrel{\text{def}}{=} Ker(B)$$

con lo que en principio se tendrá que  $\dim(W_1) = m$ .

**Proposición 2.1.8** *La restricción de la aplicación lineal  $B$  al subespacio  $W_1$ , que denotaremos con*

$$\bar{B} : W_1 \rightarrow \mathbb{R}^m$$

*resulta ser un isomorfismo lineal.*

*Demostración*

En primer lugar, de las hipótesis anteriormente consideradas se deduce que  $\bar{B}$  es una aplicación inyectiva. Para ello, téngase presente que

$$Ker(\bar{B}) = Ker(L) \cap Ker(B) = \{0\}$$

Así mismo, debido a que la aplicación lineal e inyectiva  $\bar{B}$  está definida entre dos espacios vectoriales de dimensión finita  $m$ , puede concluirse que dicha aplicación ha de ser un isomorfismo lineal.



□

El siguiente resultado muestra cómo, con las condiciones exigidas, estamos generalizando la condición de existencia y unicidad de solución para el PVI en el caso de que el operador  $L$  sea diferencial.

**Lema 2.1.2**  $\forall \alpha \in \mathbb{R}^m$  y  $\forall u \in S$  existe un único elemento  $x \in W$  cumpliendo:

$$\begin{cases} L(x) = u \\ B(x) = \alpha \end{cases}$$

*Demostración*

En principio, la unicidad es consecuencia inmediata de que

$$Ker(L) \cap Ker(B) = \{0\}$$

Por otro lado, dado  $\alpha \in \mathbb{R}^m$  y  $u \in S$  arbitrarios pero fijo, se sabe que por ser  $L$  un operador sobreyectivo existirá  $x_o \in W$  cumpliendo que  $L(x_o) = u$ . De donde puede afirmarse que  $L^{-1}(u) = x_o + Ker(L)$ . Con esto, nuestro problema consistiría en buscar  $v \in Ker(L)$  tal que  $x_o + v$  sea un elemento cumpliendo

$$\alpha = B(x_o + v) = B(x_o) + \bar{B}(v)$$

es decir, el problema consistiría en encontrar  $v \in W_1$  verificando

$$\bar{B}(v) = \alpha - B(x_o)$$

de donde, teniendo en cuenta el Lema anterior, queda garantizada la existencia de dicho elemento  $v$ .

□

**Proposición 2.1.9** *La restricción del operador lineal  $L$  al subespacio  $W_2$ , que denotaremos con*

$$\bar{L} : W_2 \longrightarrow S$$

*es un isomorfismo lineal.*

*Demostración*

(i) En relación a la inyectividad de  $\bar{L}$ , se deduce por hipótesis lo siguiente:

$$Ker(\bar{L}) = Ker(B) \cap Ker(L) = \{0\}$$

(ii) Por otro lado, la sobreyectividad es consecuencia del Lema anterior, sin más que considerar el caso de que  $\alpha = 0$ .

□

**Proposición 2.1.10** *Los subespacios vectoriales  $W_1$  y  $W_2$  cumplen:*

1.  $W_1 \cap W_2 = \{0\}$
2.  $W = W_1 \oplus W_2$
3.  $\forall f \in W \quad \begin{cases} \exists f_1 \in W_1 & (B(f_1) = B(f)) \\ \exists f_2 \in W_2 & (L(f_2) = L(f)) \end{cases} \quad \text{tal que } f = f_1 + f_2$
4.  $\dim(W_1) = m \quad \text{y} \quad \dim(W_2) = \infty$

*Demostración*

Obsérvese que la primera afirmación resulta ser una de las hipótesis de partida. Así mismo, se sabe que  $W_1 + W_2 \subseteq W$ .

Por otra parte, dado  $f \in W$ , denotemos con  $u \stackrel{\text{def}}{=} L(f) \in S$ . En principio, por la Proposición (2.1.9) puede afirmarse que:

$$\exists f_2 \in W_2 \text{ tal que } \bar{L}(f_2) = L(f_2) = u = L(f)$$

De lo anterior, si consideramos  $f_1 = f - f_2$  puede comprobarse que

$$L(f_1) = L(f) - L(f_2) = 0 \implies f_1 \in W_1$$

Por tanto, obtenemos la siguiente descomposición:

$$f = (f - f_2) + f_2 = f_1 + f_2 \in W_1 + W_2$$

□

A partir de ahora, para cada  $l \in \{1, 2\}$ , denotaremos con

$$P_l : W \longrightarrow W_l$$

a la proyección sobre  $W_l$ , de forma que:  $f = P_1(f) + P_2(f) \quad \forall f \in W$

A continuación, definiremos sendas estructuras geométricas sobre los subespacios vectoriales  $W_1$  y  $W_2$ . Una forma de llevarlo a cabo sería servirse de los isomorfismos lineales,  $\bar{B}$  y  $\bar{L}$ , cuyos codominios son respectivamente los espacios de Hilbert  $\mathbb{R}^m$  y  $S$ .

En concreto, definiríamos las siguientes aplicaciones:

$$\langle *, * \rangle_1: W_1 \times W_1 \longrightarrow \mathbb{R}_o^+$$

$$\langle f_1, g_1 \rangle_1 \stackrel{\text{def}}{=} \langle \bar{B}(f_1), \bar{B}(g_1) \rangle_u = \sum_{j=1}^m B_j(f_1) B_j(g_1) \quad \forall f_1, g_1 \in W_1$$

( notando con  $\langle *, * \rangle_u$  al producto escalar usual de  $\mathbb{R}^m$  )

$$\langle *, * \rangle_2: W_2 \times W_2 \longrightarrow \mathbb{R}_o^+$$

$$\langle f_2, g_2 \rangle_2 \stackrel{\text{def}}{=} \langle \bar{L}(f_2), \bar{L}(g_2) \rangle_s \quad \forall f_2, g_2 \in W_2$$

**Proposición 2.1.11**

1.  $\langle *, * \rangle_l$  es un producto escalar sobre  $W_l$ , para cada  $l \in \{1, 2\}$ .
2.  $\bar{B}$  es una isometría lineal, con lo que  $(W_1, \langle *, * \rangle_1) \simeq \mathbb{R}^m$
3.  $\bar{L}$  es una isometría lineal, con lo que  $(W_2, \langle *, * \rangle_2) \simeq S$
4.  $(W_l, \langle *, * \rangle_l)$  es un espacio de Hilbert, para cada  $l \in \{1, 2\}$ .

A continuación, intentaremos dotar al espacio  $W$  de estructura geométrica de forma que contenga, de alguna forma, a las dos estructuras geométricas definidas anteriormente sobre los subespacios vectoriales  $W_1$  y  $W_2$ . Una forma de llevarlo a cabo podría ser la siguiente:

$$\langle *, * \rangle_o: W \times W \longrightarrow \mathbb{R}_o^+$$

$$\langle f, g \rangle_o = \langle P_1(f), P_1(g) \rangle_1 + \langle P_2(f), P_2(g) \rangle_2 \quad \forall f, g \in W$$

**Proposición 2.1.12**

1.  $\langle *, * \rangle_o$  es un producto escalar sobre  $W$ .
2.  $(W_l, \langle *, * \rangle_o) = (W_l, \langle *, * \rangle_l)$ , para cada  $l \in \{1, 2\}$ .
3.  $W_1 \perp W_2$  en  $(W, \langle *, * \rangle_o)$ .  
En concreto,  $(W, \langle *, * \rangle_o) = (W_1, \langle *, * \rangle_1) \otimes (W_2, \langle *, * \rangle_2)$
4.  $(W, \langle *, * \rangle_o)$  posee estructura de espacio de Hilbert.

5.  $P_l$  es la proyección ortogonal sobre  $W_l$ , para cada  $l \in \{1, 2\}$ .

**Proposición 2.1.13**

$$\langle f, g \rangle_o = \langle B(f), B(g) \rangle_u + \langle L(f), L(g) \rangle_s \quad \forall f, g \in G$$

*Demostración*

Consideremos  $f, g \in W$  arbitrarios pero fijos, y denotemos sus correspondientes descomposiciones:  $f = f_1 + f_2$  y  $g = g_1 + g_2$ ; siendo, para cada  $l \in \{1, 2\}$ ,  $f_l = P_l(f)$  y  $g_l = P_l(g)$ .

En estas condiciones, se razona como sigue:

$$\begin{aligned} \langle B(f), B(g) \rangle_u + \langle L(f), L(g) \rangle_s &= \langle B(f_1), B(g_1) \rangle_u + \\ &+ \langle L(f_2), L(g_2) \rangle_s = \langle f_1, g_1 \rangle_1 + \langle f_2, g_2 \rangle_2 = \langle f, g \rangle_o \end{aligned}$$

□

Por otra parte, podemos considerar el espacio de Hilbert  $\mathbb{R}^m \times S$  con el producto escalar definido por

$$\langle (v, s_1), (w, s_2) \rangle_x \stackrel{\text{def}}{=} \langle v, w \rangle_u + \langle s_1, s_2 \rangle_s \quad \forall (v, s_1), (w, s_2) \in \mathbb{R}^m \times S$$

**Corolario 2.1.3** *La aplicación*

$$F : (W, \langle *, * \rangle_o) \mapsto (\mathbb{R}^m \times S, \langle *, * \rangle_x)$$

$$F(f) \stackrel{\text{def}}{=} (B(f), L(f)) \quad \forall f \in W$$

es una isometría lineal.

*Demostración*

La demostración es consecuencia inmediata de la anterior Proposición.

□

De esta forma, queda justificado la metodología en la situación expuesta en la introducción para el caso particular del espacio funcional de outputs  $W^m$  y para un operador diferencial, en la que se tiene que  $S = L_2(T)$ . Es decir, considerando la notación allí introducida:

- El ACP de la familia de outputs  $\{x_i : i = 1, \dots, n\}$  con respecto a la geometría  $(W^m, \langle *, * \rangle_o)$  es equivalente al ACP de  $\{F(x_i) : i = 1, \dots, n\}$  con respecto a la geometría  $(\mathbb{R}^m \times L_2(T), \langle *, * \rangle_x)$ .
- El ACP de la familia de funciones  $\{x_i^{(1)} : i = 1, \dots, n\}$  con respecto a la geometría  $(W_1^m, \langle *, * \rangle_1)$  es equivalente al ACP de  $\{B(x_i) : i = 1, \dots, n\}$  con respecto a la geometría usual de  $\mathbb{R}^m$ .
- El ACP de la familia de funciones  $\{x_i^{(2)} : i = 1, \dots, n\}$  con respecto a la geometría  $(W_2^m, \langle *, * \rangle_2)$  es equivalente al ACP de  $\{L(x_i) = u_i : i = 1, \dots, n\}$  con respecto a la geometría de  $L_2(T)$ .

Téngase presente que en este caso se obtendría:

$$x_i^{(l)} = P_l(x_i) \quad \forall i \in \{1, \dots, n\} \quad \forall l \in \{1, 2\}$$

## §2.2 ESTUDIO SOBRE LA MODIFICACION DE LA GEOMETRIA

En esta sección consideraremos un espacio de Hilbert  $(H, \tau)$ , y

$$X : (\Omega, \mathcal{A}, P) \mapsto (H, \mathcal{B}_\tau) \quad \text{una aplicación } \mathcal{B}_\tau/\mathcal{A} \text{ medible.}$$

Denotaremos con  $\mathcal{B}_\tau$  a la  $\sigma$ -álgebra generada por la topología asociada al producto escalar  $\tau$ , que determinará la estructura geométrica *inicial de referencia*.

En esta situación, podría estudiarse la influencia sobre  $X$  de una modificación en la geometría de  $H$ , definida inicialmente por el producto escalar  $\tau$ . Es decir, analizaremos el comportamiento de algunas características de la v. a.  $X$  frente a otro producto escalar.

### 2.2.1 CONSIDERACIONES SOBRE LA CLASE DE PRODUCTOS ESCALARES EN ESTUDIO

En esta sección se pretende: analizar algunas propiedades que *convendría* exigir a la clase de productos escalares sobre  $H$ ; así como estudiar la relación que pueda existir con el producto escalar inicialmente considerado  $\tau$ .

En principio, una modificación *geométrica* a considerar no debiera alterar la condición de medibilidad de  $X$ . Una forma de conseguirlo, sería intentar buscar productos escalares sobre  $H$ , de forma que, la topología generada por ellos esté contenida en la asociada a  $\tau$ . Con esto, se tendría la misma relación de contenido para las correspondientes  $\sigma$ -álgebras de Borel asociadas: manteniendo así invariante la propiedad de *medibilidad* de  $X$ .

**Proposición 2.2.1** *Supóngase que  $(H, \tau)$  fuese un espacio prehilbertiano, y*

$$\rho : H \times H \mapsto \mathbb{R} \quad \text{un producto escalar sobre } H.$$

*Entonces, son equivalentes las siguientes afirmaciones:*

1.  $T_\rho \subseteq T_\tau$
2.  $\rho$  es continua sobre el espacio topológico  $(H \times H, T_\tau \otimes T_\tau)$
3.  $\exists K > 0$  tal que  $\rho(f, f) \leq K^2 \tau(f, f) \quad \forall f \in H$

Notaremos con  $T_\rho$  y  $T_\tau$  a las topologías definidas respectivamente por los productos escalares  $\rho$  y  $\tau$ .

*Demostración*

(1)  $\implies$  (2)

En principio, sabemos que el producto escalar  $\rho$  será una aplicación continua sobre el espacio topológico  $(H \times H, T_\rho \otimes T_\rho)$ .

Por tanto, teniendo en cuenta lo anterior y el hecho de que por hipótesis se verificaría que  $T_\rho \otimes T_\rho \subseteq T_\tau \otimes T_\tau$ , podemos concluir que la aplicación  $\rho : (H \times H, T_\tau \otimes T_\tau) \rightarrow \mathbb{R}$  es continua.

(2)  $\implies$  (3)

Haciendo uso de una de las caracterizaciones de la continuidad de aplicaciones bilineales, puede afirmarse que existe  $K > 0$  tal que

$$|\rho(f, g)| \leq K^2 \|f\|_\tau \|g\|_\tau \quad \forall f, g \in H$$

con lo que  $\rho(f, f) \leq K^2 \tau(f, f) \quad \forall f \in H$

(3)  $\implies$  (1)

Puede observarse que para cualquier bola abierta de  $T_\rho$

$$B_\rho(r) \stackrel{\text{def}}{=} \{f \in H : \rho(f, f) < r^2\} \quad \text{con } r \in \mathbb{R}^+$$

podemos considerar la bola abierta de  $T_\tau$  dada por

$$B_\tau(K^{-1}r) \stackrel{\text{def}}{=} \{f \in H : \tau(f, f) < K^{-2}r^2\}$$

la cuál, puede comprobarse que verifica:  $B_\tau(K^{-1}r) \subseteq B_\rho(r)$ .

Por tanto, haciendo uso de que el espacio vectorial  $H$  con ambas topologías tiene estructura de espacio vectorial topológico, puede concluirse que

$$T_\rho \subseteq T_\tau.$$

□

En definitiva, la propiedad de medibilidad de  $X$ , dada por la condición inicial:

$$X^{-1}(G) \in \mathcal{A} \quad \forall G \in T_\tau$$

se mantendrá bajo la hipótesis de continuidad de  $\rho$ , según indica la Proposición anterior, pues evidentemente se cumplirá:

$$X^{-1}(G') \in \mathcal{A} \quad \forall G' \in T_\rho$$

Con el resultado anterior no se pretende justificar que la clase de productos escalares, con los que la v. a. hilbertiana  $X$  sigue siendo medible, tengan que ser todos continuos con respecto a la topología  $T_\tau$ . Si no que, entre ellos, hemos garantizado que se encontrarán los que cumplan esa propiedad de continuidad.

Téngase en cuenta que pretendemos analizar los productos escalares que, además de *conservar* la propiedad de medibilidad de  $X$ , puedan obtenerse a partir del producto escalar de referencia  $\tau$  mediante ciertas transformaciones continuas, en un cierto sentido que más adelante se especificará. Por esta razón, a partir de ahora consideraremos un producto escalar continuo

$$\rho : (H \times H, T_\tau \otimes T_\tau) \mapsto \mathbb{R}$$

En el siguiente ejemplo se muestran diferentes productos escalares cumpliendo la propiedad de continuidad objeto de nuestro interés.

**Ejemplo 2.2.1** *Sobre un espacio de Hilbert  $(H, \tau)$  se consideran los productos escalares siguientes:*

$$\rho_1(f, g) = \tau(D(f), D(g)) \quad \forall f, g \in H$$

$$\rho_2(f, g) = \tau(S(f), g) \quad \forall f, g \in H$$

*serían continuos en el espacio topológico  $(H \times H, T_\tau \otimes T_\tau)$*

*Donde con  $D$  y  $S$  estamos denotando a operadores lineales, inyectivos y continuos con respecto a  $T_\tau$ , y siendo además  $S$  un operador simétrico y definido positivo.*

En el siguiente Teorema caracterizamos la forma que pueden adoptar los productos escalares sobre  $H$ , cuando además son continuos en el espacio topológico  $(H \times H, T_\tau \otimes T_\tau)$ .

**Teorema 2.2.1** *Sean  $(H, \tau)$  un espacio de Hilbert, y un producto escalar,  $\rho : H \times H \mapsto \mathbb{R}$ , continuo con la topología  $T_\tau \otimes T_\tau$ .*

*Entonces, existe un único operador,  $S : H \mapsto H$ , cumpliendo:*

$$\rho(f, g) = \tau(S(f), g) \quad \forall f, g \in H \quad (2.1)$$



*Demostración*

En principio, probaremos la existencia del operador  $S : H \mapsto H$ . Obsérvese que la unicidad se deduce fácilmente a partir de la propiedad que lo caracteriza.

Dado  $f \in H$  arbitrario pero fijo, consideraremos la aplicación

$$\begin{aligned} \gamma_f : (H, \tau) &\mapsto \mathbb{R} \\ \gamma_f(g) &:= \rho(f, g) \quad \forall g \in H \end{aligned}$$

Por ser  $\rho$  una aplicación bilineal y continua, con respecto al espacio topológico  $(H \times H, T_\tau \otimes T_\tau)$ , puede afirmarse que

$$\gamma_f \in (H, \tau)'$$

Haciendo uso del Teorema de Riesz-Frechet, puede garantizarse la existencia de un *único* elemento de  $H$ , que denotaremos con  $S(f)$ , tal que

$$\gamma_f(g) = \tau(S(f), g) \quad \forall g \in H$$

cumpléndose además que

$$\|S(f)\|_\tau = \|\gamma_f\|_{(H, \tau)'} = \sup_{g \in S_H^\tau} \gamma_f(g) \tag{2.2}$$

□

Al hilo del anterior Teorema, en las dos siguientes Proposiciones continuaremos analizando algunas propiedades del operador  $S$ , asociado al producto escalar  $\rho$  bajo la hipótesis de continuidad anteriormente establecida.

**Proposición 2.2.2** *En las condiciones del anterior Teorema, el operador*

$$S : (H, \tau) \mapsto (H, \tau)$$

*es lineal, y para el producto escalar  $\tau$ : continuo, simétrico y definido positivo. Cumpléndose además que*

$$\sup_{f \in S_H^\tau} \rho(f, f) = \|S\|_{BL}$$

Donde estamos denotando con:

$$S_H^\tau \stackrel{\text{def}}{=} \{f \in H : \|f\|_\tau = 1\}$$

$$BL(H, \tau) \stackrel{\text{def}}{=} \{F : (H, T_\tau) \mapsto (H, T_\tau) / F \text{ es lineal y continuo}\}$$

$$\|F\|_{BL} \stackrel{\text{def}}{=} \sup_{f \in S_H^\tau} \|F(f)\|_\tau \quad \forall F \in BL(H, \tau)$$

*Demostración*

1. *S es lineal.*

Dados  $f_1, f_2 \in H$  y  $\alpha, \beta \in \mathbb{R}$ , se tiene que  $\forall g \in H$

$$\begin{aligned} \tau(S(\alpha f_1 + \beta f_2), g) &= \rho(\alpha f_1 + \beta f_2, g) = \alpha \rho(f_1, g) + \beta \rho(f_2, g) = \\ &= \alpha \tau(S(f_1), g) + \beta \tau(S(f_2), g) = \tau(\alpha S(f_1) + \beta S(f_2), g) \end{aligned}$$

de donde concluimos que  $S(\alpha f_1 + \beta f_2) = \alpha S(f_1) + \beta S(f_2)$

2. *S es un operador simétrico para  $\tau$ .*

Dados  $f, g \in H$ , se tiene

$$\tau(S(f), g) = \rho(f, g) = \rho(g, f) = \tau(S(g), f) = \tau(f, S(g))$$

3. *S es un operador definido positivo, con respecto a  $\tau$ , por ser  $\rho$  definido positivo.*

4.  *$S \in BL(H, \tau)$*

Por ser  $\rho$  una aplicación bilineal y continua, con respecto al producto escalar  $\tau$ , se puede afirmar que existe el siguiente supremo

$$M \stackrel{\text{def}}{=} \sup_{f, g \in S_H^\tau} |\rho(f, g)| = \sup_{f, g \in S_H^\tau} \rho(f, g)$$

Con lo que haciendo uso de lo obtenido en (2.2), para cada  $f \in H - \{0\}$ , se cumple:

$$\|S(f)\|_\tau = \sup_{g \in S_H^\tau} |\rho(f, g)| = \|f\|_\tau \sup_{g \in S_H^\tau} \rho\left(\frac{f}{\|f\|_\tau}, g\right) \leq \|f\|_\tau M$$

Por tanto, concluimos que  $S \in BL(H, \tau)$ , y además  $\|S\|_{BL} \leq M$

Por otra parte, haciendo uso de que  $S$  es un operador lineal simétrico, continuo y definido positivo se tiene:

$$\|S\|_{BL} = \sup_{f \in S_H^\tau} \tau(S(f), f) = \sup_{f \in S_H^\tau} \rho(f, f)$$

□

Recíprocamente, una aplicación sobre  $H \times H$ , definida como en (2.1) a partir de un operador cumpliendo las propiedades de la Proposición anterior, resulta ser un producto escalar *continuo*. En esencia, con la Proposición anterior hemos caracterizado la clase de productos escalares continuos, con respecto al producto escalar  $\tau$ .

A continuación afinaremos en el análisis de la continuidad del operador  $S$ .

**Proposición 2.2.3** *En las condiciones del Teorema anterior, el operador*

$$S : (H, \rho) \mapsto (H, \tau) \quad \text{es continuo.}$$

**Observación:** La consideración en el dominio de  $S$  de la topología  $T_\rho$ , que sabemos cumple  $T_\rho \subseteq T_\tau$ , implica un afinamiento en la propiedad de continuidad establecida en la anterior Proposición.

*Demostración*

Dado  $h \in H$  arbitrario pero fijo, obsérvese que utilizando la desigualdad de Cauchy-Schwarz se obtiene:

$$\|S(h)\|_\tau^2 = \tau(S(h), S(h)) = \rho(h, S(h)) \leq \|h\|_\rho \|S(h)\|_\rho$$

Por otro lado, sabemos que se cumple  $\rho(h, h) \leq K^2 \tau(h, h) \quad \forall h \in H$  por lo que se verificará que  $\|S(h)\|_\rho \leq K \|S(h)\|_\tau$ .

Uniendo los dos anteriores razonamientos se obtiene que

$$\|S(h)\|_\tau^2 \leq \|h\|_\rho K \|S(h)\|_\tau$$

de donde se concluye que  $\|S(h)\|_\tau \leq K \|h\|_\rho$ .

□

Consideremos los siguientes conjuntos

$$\text{SimBL}(H, \tau) \stackrel{\text{def}}{=} \{S \in \text{BL}(H, \tau) : S \text{ es simétrico}\}$$

$$\text{Pos}(H, \tau) \stackrel{\text{def}}{=} \{S \in \text{SimBL}(H, \tau) : S \text{ es definido positivo}\}$$

Como es fácil comprobar, el conjunto  $\text{SimBL}(H, \tau)$  tiene estructura de espacio vectorial, que contiene al cono convexo  $\text{Pos}(H)$ . Además, la clase de productos escalares sobre  $H$  continuos con respecto a  $(H, \tau)$  es biyectivamente idéntica a  $\text{Pos}(H, \tau)$ . En el siguiente resultado, se analizan algunas propiedades de los operadores del conjunto  $\text{Pos}(H, \tau)$ .

**Proposición 2.2.4** *Sea  $S \in \text{Pos}(H, \tau)$ , entonces se cumple:*

1.  $S$  es inyectivo.
2.  $\overline{S(H)}^\tau = H$
3.  $\|S\|_{BL} = \sup_{f \in \overline{S(H)}^\tau} \rho(f, f)$
4. Existe el operador  $S^{-1} : (S(H), \tau) \mapsto (H, \tau)$ , que es lineal, autoadjunto, y definido positivo.

*Demostración*

1.  $S$  es inyectivo.

Dado  $f \in \text{Ker}(S)$ , se tiene  $\tau(S(f), f) = 0$ . Por ser,  $S$  definido positivo, se concluye que  $f = 0$ . De donde se deduce que  $\text{Ker}(S) = \{0\}$

2.  $S(H)$  es denso en el espacio de Hilbert  $(H, \tau)$ .

Consideremos el subespacio vectorial  $N \stackrel{\text{def}}{=} \overline{S(H)}^\tau$ . Como  $N$  es un subespacio vectorial cerrado en  $(H, \tau)$ , dicho espacio de Hilbert podrá descomponerse de la forma siguiente:

$$H = N \oplus N^\perp$$

siendo  $N^\perp \stackrel{\text{def}}{=} \{f \in H : \tau(f, h) = 0 \ \forall h \in N\}$

Teniendo en cuenta esto, dado  $f \in N^\perp$ , se tendría que

$$\tau(f, S(f)) = 0.$$

Con lo que por ser  $S$  un operador definido positivo, se obtiene que  $f = 0$ . En definitiva, puede afirmarse que  $N^\perp = \{0\}$ , y concluir así que  $H = N$

3. *Sobre la existencia de  $S^{-1}$*

En principio, por lo ya analizado, es obvio que existirá el operador

$$S^{-1} : S(H) \mapsto H$$

heredando, en principio, la propiedad de linealidad de  $S$ .

Probemos que el operador  $S^{-1}$  es autoadjunto. Para ello, analizaremos el dominio del adjunto del operador  $S^{-1}$ .

En principio, se sabe que  $h \in \text{Dom}(S^{-1})^*$ , si y solamente si,

$$\exists h' \in H \quad \text{tal que } \tau(h, S^{-1}(g)) = \tau(h', g) \quad \forall g \in S(H)$$

o equivalentemente  $\tau(h, f) = \tau(h', S(f)) \quad \forall f \in H$

Esta expresión, a su vez, teniendo en cuenta que  $S$  es simétrico, podría ser reformulada de la siguiente manera

$$\tau(h, f) = \tau(h', S(f)) = \tau(S(h'), f) \quad \forall f \in H$$

Del razonamiento anterior, obtenemos que:  $h \in \text{Dom}(S^{-1})^*$ , si y solamente si,

$$\exists h' \in H \text{ tal que } S(h') = h \implies h' = S^{-1}(h)$$

En definitiva, de lo anterior se deduce:

(a)  $\text{Dom}(S^{-1})^* = S(H)$

(b)  $(S^{-1})^* = S^{-1}$

Por otra parte, para cualquier  $g = S(f)$ , con  $f \in H$ , se tiene

$$\tau(S^{-1}(g), g) = \tau(f, S(f))$$

de donde, puede deducirse que el operador  $S^{-1}$  es definido positivo.

□

### 2.2.2 ACP DE UNA V.A. HILBERTIANA FRENTE A UNA MODIFICACION EN LA GEOMETRIA

Como siempre, consideremos un espacio de Hilbert  $(H, \tau)$ , como marco de referencia, y  $X : (\Omega, \mathcal{A}, P) \mapsto (H, \mathcal{B}_\tau)$  una aplicación  $\mathcal{B}_\tau/\mathcal{A}$  medible.

Así mismo, consideremos un producto escalar sobre  $H$

$$\rho : H \times H \mapsto \mathbb{R}$$

continuo sobre la topología  $T_\tau \otimes T_\tau$ . Y notemos con  $S \in \text{Pos}(H, \tau)$  a su operador asociado, es decir:

$$\rho(f, g) = \tau(S(f), g) \quad \forall f, g \in H$$

En estas condiciones, podría considerarse el espacio prehilbertiano  $(H, \rho)$ , con el que la aplicación  $X$ , por lo ya analizado en la sección anterior,

$$X : (\Omega, \mathcal{A}, P) \mapsto (H, \mathcal{B}_\rho) \quad \text{es } \mathcal{B}_\rho/\mathcal{A} \text{ medible.}$$

A partir de ahora, utilizaremos la siguiente notación, para cada  $p \in \mathbb{N}$ , y  $\varrho$  un producto escalar sobre  $H$ , se define:

$$L_p(\Omega, \mathcal{A}, P; H, \varrho) \stackrel{\text{def}}{=} \left\{ X : \Omega \mapsto (H, \varrho) \left/ \begin{array}{l} X \text{ es } \mathcal{B}_\varrho/\mathcal{A} \text{ medible} \\ \int_\Omega \|X(w)\|_\varrho^p dP(w) < \infty \end{array} \right. \right\}$$

#### Proposición 2.2.5

Si  $X \in L_p(\Omega, \mathcal{A}, P; H, \tau)$  entonces  $X \in L_p(\Omega, \mathcal{A}, P; H, \rho)$ .

*Demostración*

Obsérvese que  $\forall w \in \Omega$  se tiene

$$\begin{aligned} \|X(w)\|_\rho^p &= \rho(X(w), X(w))^{\frac{p}{2}} = \tau(S(X(w)), X(w))^{\frac{p}{2}} \leq \\ &\leq \|S(X(w))\|_\tau^{\frac{p}{2}} \|X(w)\|_\tau^{\frac{p}{2}} \leq \|S\|_{BL}^{\frac{p}{2}} \|X(w)\|_\tau^p \end{aligned}$$

De lo anterior, como la v. a. real  $\|X\|_\tau^p$  es integrable, se tiene entonces que  $\|X\|_\rho^p$  también lo es. Por tanto, se obtiene que  $X \in L_p(\Omega, \mathcal{A}, P; H, \rho)$ .

□

**Proposición 2.2.6** *Si existe la esperanza de  $X$  con respecto al producto escalar  $\tau$ , entonces también existe la esperanza con respecto al producto escalar  $\rho$ .*

*Además, en caso afirmativo se cumple*

$$E^\rho[X] = E^\tau[X]$$

*Demostración*

Téngase presente que buscamos un elemento  $E^\rho[X] \in H$ , cumpliendo:

$$\rho(E^\rho[X], f) = \int_{\Omega} \rho(X(w), f) dP(w) \quad \forall f \in H$$

Por otro lado,  $\forall f \in H$ , puede razonarse del siguiente modo

$$\begin{aligned} \int_{\Omega} \rho(X(w), f) dP(w) &= \int_{\Omega} \tau(X(w), S(f)) dP(w) = \\ &= \tau(E^\tau[X], S(f)) = \rho(E^\tau[X], f) \end{aligned}$$

concluyéndose así que  $E^\rho[X] = E^\tau[X]$ .

□

De esta forma, podemos afirmar que la esperanza matemática es invariante frente a cambios en la geometría del espacio considerado, dentro de la clase de productos escalares *continuos* en estudio. Por esta razón, a partir de ahora la denotaremos con  $\mathbf{E}[X]$ , sin especificar el producto escalar considerado.

**Proposición 2.2.7** *Si  $X \in L_2(\Omega, \mathcal{A}, P; H, \tau)$  se cumple:*

1.  $X \in L_2(\Omega, \mathcal{A}, P; H, \rho)$
2.  $U_x^\rho = U_x^\tau \circ S$
3.  $U_x^{\rho*} = U_x^{\tau*}$
4.  $C_x^\rho = C_x^\tau \circ S$
5.  $W_x^\rho = U_x^\tau \circ S \circ U_x^{\tau*}$

*Demostración*

A partir de ahora, denotaremos con  $\tilde{X}(w) \stackrel{\text{def}}{=} X(w) - \mathbf{E}[X] \quad \forall w \in \Omega$

Consideremos el operador  $U_x^\rho : (H, \rho) \mapsto L_2(\Omega, \mathcal{A}, P)$  definido, para cada  $f \in H$ , como sigue:

$$U_x^\rho(f)(w) \stackrel{\text{def}}{=} \rho(f, \tilde{X}(w)) = \tau(S(f), \tilde{X}(w)) = U_x^\tau(S(f))(w) \quad \forall w \in \Omega$$

de donde se concluye que  $U_x^\rho = U_x^\tau \circ S$ .

A continuación estudiaremos la relación existente entre los operadores adjuntos de  $U_x^\rho$  y  $U_x^\tau$ , cada uno de ellos con respecto al *correspondiente* producto escalar. En concreto, consideremos el operador  $U_x^{\rho*}$  definido por la propiedad siguiente:

$$\langle U_x^\rho(f), \psi \rangle_2 = \rho(f, U_x^{\rho*}(\psi)) \quad \forall f \in H \quad \forall \psi \in L_2(\Omega, \mathcal{A}, P)$$

Por otro lado, obsérvese que

$$\langle U_x^\rho(f), \psi \rangle_2 = \langle U_x^\tau(S(f)), \psi \rangle_2 = \tau(S(f), U_x^{\tau*}(\psi)) = \rho(f, U_x^{\tau*}(\psi))$$

de donde se concluye que  $U_x^{\rho*} = U_x^{\tau*}$ .

□

**Teorema 2.2.2** *Supóngase que existen  $(G, \langle *, * \rangle_G)$  un espacio de Hilbert, y  $F : (H, \tau) \mapsto (G, \langle *, * \rangle_G)$  un operador lineal, continuo y no singular, tal que  $S = F^* \circ F$ . Además, supondremos que  $(H, \rho)$  posee estructura de espacio de Hilbert.*

*Entonces, una representación en componentes principales de la v.a.  $F \circ X$  con respecto a la geometría del espacio  $(G, \langle *, * \rangle_G)$ , dada por*

$$F \circ X(w) = \sum_{i \in I} F(f_i) \psi_i(w),$$

*determina una representación en componentes principales de la v.a.  $X$  con respecto a la geometría inducida por el nuevo producto escalar  $\rho$  de la forma siguiente:*

$$X(w) = \sum_{i \in I} f_i \psi_i(w)$$

donde  $\forall i \in I$

$$C_x^\rho(f_i) = \lambda_i f_i \quad \psi_i \stackrel{\text{def}}{=} U_x^\rho(f_i)$$



*Demostración*

Obsérvese que, en estas condiciones, podríamos considerar al operador lineal  $F$  definido entre los espacios

$$F : (H, \rho) \mapsto F(H) \subseteq (G, \langle *, * \rangle_G)$$

cumpliéndose que

$$\langle F(h_1), F(h_2) \rangle_G = \tau(F^* \circ F(h_1), h_2) = \tau(S(h_1), h_2) = \rho(h_1, h_2)$$

Con lo que podemos concluir que  $F : (H, \rho) \mapsto (F(H), \langle *, * \rangle_G)$  es una isometría lineal.

Teniendo en cuenta este resultado, estamos en las condiciones del Teorema que resuelve el problema del ACP filtrante.

□

El enunciado anterior ha sido enunciado en un ambiente lo más general posible. Sin embargo, téngase en cuenta que, por cumplir el operador  $S$  las propiedades enunciadas en la Proposición (2.2.4), queda asegurada la existencia del operador (*raiz cuadrada*)

$$F : (H, \tau) \mapsto (H, \tau)$$

cumpliendo, además de las mismas propiedades anteriormente citadas, que  $S = F \circ F$ .

### §2.3 ACP DE UNA V.A. HILBERTIANA CON VALORES EN UN ESPACIO DE DIMENSION FINITA

En esta sección, como aplicación de lo obtenido en las secciones anteriores, analizaremos el problema de la realización del ACP de una v.a. hilbertiana con valores en un espacio finito dimensional. La importancia de este tipo de v.a. radica en el hecho de que en problemas de aproximación de ACP de una v.a. hilbertiana, existen diversas metodologías en las que la forma de resolver dicho problema consiste en realizar el ACP de una cierta v.a. hilbertiana con valores en un cierto espacio de dimensión finita, la cuál se propone como *aproximación* de la correspondiente v.a. en estudio.

Sea  $X : (\Omega, \mathcal{A}, P) \mapsto \mathbb{E}$  una v. a. hilbertiana de  $L_2(\Omega, \mathcal{A}, P; \mathbb{E})$ , siendo  $\mathbb{E}$  un espacio vectorial de dimensión  $m$  sobre el que tenemos definido un producto escalar  $\langle *, * \rangle_E$ .

En principio, para una cierta base del espacio  $\mathbb{E}$ , que denotaremos con  $\mathcal{B} = \{e_1, \dots, e_m\}$ , la v. a.  $X$  admite, para cada  $w \in \Omega$ , la representación lineal siguiente:

$$X(w) = \sum_{i=1}^m Y_i(w) e_i \quad \forall w \in \Omega \quad (2.1)$$

De esta forma, obtenemos una colección de funciones

$$\{Y_i : \Omega \mapsto \mathbb{R} : i = 1, \dots, m\}$$

que determinan a  $X$  a partir de la base  $\mathcal{B}$ , a las cuales daremos el calificativo de *coordenadas*, y con las que definiremos la función vectorial  $Y \stackrel{\text{def}}{=} (Y_1, \dots, Y_m)^t$ .

**Proposición 2.3.1** *En las condiciones anteriores, se cumple que*

$$Y_i \in L_2(\Omega, \mathcal{A}, P) \quad \forall i = 1, \dots, m$$

*Demostración*

En principio, consideraremos, para cada  $k \in \{1, \dots, m\}$ , la v.a. de  $L_2(\Omega, \mathcal{A}, P)$  siguiente:

$$Z_k(w) \stackrel{\text{def}}{=} u_x(e_k) = \langle e_k, X(w) \rangle_E \quad \forall w \in \Omega$$

y notemos con  $Z \stackrel{\text{def}}{=} (Z_1, \dots, Z_m)^t \in L_2(\Omega, \mathcal{A}, P; \mathbb{R}^m)$ .

Obsérvese que de la expresión (2.1) se obtiene que

$$Z_k(w) = \sum_{i=1}^m Y_i(w) \langle e_k, e_i \rangle_E \quad \forall w \in \Omega$$

equivalentemente en forma matricial podría expresarse:

$$Z(w) = \mathbf{P} Y(w) \quad \forall w \in \Omega$$

donde estamos denotando con

$$\mathbf{P} \stackrel{\text{def}}{=} (\langle e_i, e_j \rangle_E : i, j = 1, \dots, m) \in M_m(\mathbb{R})$$

Por ser la matriz cuadrada  $\mathbf{P}$  no singular, obtenemos de la expresión anterior que

$$Y = \mathbf{P}^{-1} Z$$

de donde podemos afirmar que  $Y$  es un vector aleatorio de segundo orden.

□

**Lema 2.3.1** *La matriz simétrica  $\mathbf{P}$  es definida positiva, y por tanto no singular.*

*Demostración*

En principio, por la forma de definir la matriz  $\mathbf{P}$  puede concluirse que es simétrica.

Por otra parte, dado  $v = (v_1, \dots, v_m)^t \in \mathbb{R}^m$ , se obtiene

$$v^t \mathbf{P} v = \left\| \sum_{j=1}^m v_j e_j \right\|_E^2$$

de donde se concluye que:

- $v^t \mathbf{P} v \geq 0 \quad \forall v \in \mathbb{R}^m$
- Si  $v^t \mathbf{P} v = 0$  entonces  $\sum_{j=1}^m v_j e_j = 0$ , con lo que haciendo uso de que  $\mathcal{B}$  es una base se concluye que  $v = 0$ .

□

En principio, para el vector aleatorio  $Y = (Y_1, \dots, Y_m)^t \in L_2(\Omega, \mathcal{A}, P; \mathbb{R}^m)$ , podemos considerar los siguientes elementos:

$$\mu \stackrel{\text{def}}{=} \mathbf{E}[Y] \quad C_Y \stackrel{\text{def}}{=} \mathbf{E}[(Y - \mu)(Y - \mu)^t]$$

obtenidos ambos elementos considerando la geometría usual sobre  $\mathbb{R}^m$ .

En lo sucesivo nos serán de gran interés las *coordenadas* con respecto a la base  $\mathcal{B}$  considerada. Por esta razón, consideraremos el isomorfismo lineal siguiente:

$$\gamma: \mathbb{R}^m \mapsto \mathbb{E}$$

$$\gamma(\alpha) = \sum_{i=1}^m a_i e_i \quad \forall \alpha = (a_1, \dots, a_m)^t \in \mathbb{R}^m$$

A partir de dicho isomorfismo lineal, podemos definir una geometría sobre  $\mathbb{R}^m$ , inducida por la del espacio  $\mathbb{E}$ , de la forma siguiente:

$$\langle \alpha, \beta \rangle_P \stackrel{\text{def}}{=} \langle \gamma(\alpha), \gamma(\beta) \rangle_E \quad \forall \alpha, \beta \in \mathbb{R}^m$$

Obsérvese cómo con este isomorfismo se obtiene:  $X = \gamma(Y)$ .

De esta forma, nos encontramos en un caso particular de lo que en la sección anterior hemos estudiado como ACP filtrante. En nuestro caso, el filtro viene determinado por el isomorfismo lineal  $\gamma$ , y el papel de espacio de datos filtrados lo realiza el espacio  $\mathbb{E}$ .

En la siguiente proposición recogemos algunas propiedades inmediatas, relativas a la estructura anteriormente definida sobre  $\mathbb{R}^m$ .

**Proposición 2.3.2** *Teniendo en cuenta la aplicación*

$$\gamma: (\mathbb{R}^m, \langle *, * \rangle_P) \mapsto \mathbb{E} \quad \text{se cumple:}$$

1.  $\langle *, * \rangle_P$  es un producto escalar sobre  $\mathbb{R}^m$ .
2.  $\gamma$  es una isometría lineal.
3.  $\langle \alpha, \beta \rangle_P = \alpha^t \mathbf{P} \beta \quad \forall \alpha, \beta \in \mathbb{R}^m$
4. Los espacios  $(\mathbb{R}^m, \langle *, * \rangle_P)$  y  $\mathbb{E}$  son isométricamente iguales.

5.  $\gamma^* = \gamma^{-1}$

siendo  $\gamma^*$  el operador adjunto del operador  $\gamma$  definido entre los espacios  $(\mathbb{R}^m, \langle *, * \rangle_P)$  y  $\mathbb{E}$ .

A continuación, mostraremos la relación existente entre algunas de las características de la v.a. hilbertiana  $X$  respecto a la geometría del espacio  $\mathbb{E}$ , y las correspondientes asociadas al vector aleatorio  $Y$  respecto de la geometría  $(\mathbb{R}^m, \langle *, * \rangle_P)$ . Propiedades éstas que son inmediatas a partir de las Proposiciones (2.1.5) y (2.1.6).

**Proposición 2.3.3**

1.  $\mathbf{E}[X] = \gamma(\mathbf{E}^{(P)}[Y]).$

2.  $C_X = \gamma \circ C_Y^{(P)} \circ \gamma^{-1}$

siendo  $C_Y^{(P)}$  y  $\mathbf{E}^{(P)}[Y]$ , respectivamente, el operador de covarianzas y la esperanza del vector aleatorio  $Y$  con respecto a la geometría  $(\mathbb{R}^m, \langle *, * \rangle_P)$ .

De la misma forma, particularizando lo establecido en el Teorema (2.1.1) en relación al ACP filtrante, puede enunciarse el siguiente resultado.

**Proposición 2.3.4** *El ACP de  $X$  en el espacio  $\mathbb{E}$  es equivalente al ACP del vector  $Y$  con respecto a la geometría  $(\mathbb{R}^m, \langle *, * \rangle_P)$ , en el sentido siguiente:*

$$Y(w) = \sum_{j=1}^m \psi_j(w) v_j \quad (\text{ respecto a } (\mathbb{R}^m, \langle *, * \rangle_P))$$

$$X(w) = \sum_{j=1}^m \psi_j(w) \gamma(v_j) \quad (\text{ respecto a } \mathbb{E})$$

□

A su vez, las características mencionadas en la anterior Proposición para el vector aleatorio  $Y$  respecto a la geometría  $(\mathbb{R}^m, \langle *, * \rangle_P)$ , pueden relacionarse con las obtenidas considerando la geometría usual de  $\mathbb{R}^m$ . Para ello, téngase en cuenta que

$$\langle \alpha, \beta \rangle_P = \langle \alpha, \mathbf{P} \beta \rangle_u \quad \forall \alpha, \beta \in \mathbb{R}^m$$

siendo  $\langle *, * \rangle_u$  el producto escalar usual de  $\mathbb{R}^m$ . Con lo cuál, la estructura geométrica del espacio  $(\mathbb{R}^m, \langle *, * \rangle_P)$  puede ser considerada como un caso particular de modificación ( *continua* ) de la estructura geométrica del espacio  $(\mathbb{R}^m, \langle *, * \rangle_u)$  a través del operador simétrico  $\mathbf{P}$ .

La siguiente Proposición recoge algunas particularizaciones de los resultados obtenidos en las Proposiciones (2.2.6) y (2.2.7).

### Proposición 2.3.5

$$1. \mathbf{E}^{(P)}[Y] = \mathbf{E}[Y]$$

$$2. \mathcal{C}_y^{(P)} = \mathcal{C}_y \mathbf{P}$$

□

Téngase en cuenta que por ser  $\mathbf{P}$  una matriz simétrica definida positiva puede considerarse la matriz simétrica  $\mathbf{Q} \stackrel{\text{def}}{=} \mathbf{P}^{\frac{1}{2}}$ , considerándola como aplicación lineal de  $(\mathbb{R}^m, \langle *, * \rangle_u)$  en sí mismo, con lo que  $\mathbf{Q}^* = \mathbf{Q}^t = \mathbf{Q}$ , y obteniéndose que

$$\mathbf{P} = \mathbf{Q}^* \mathbf{Q}$$

En estas condiciones, podemos particularizar lo establecido en el Teorema (2.2.2) en la siguiente Proposición.

**Proposición 2.3.6** *El ACP del vector aleatorio  $Y$  con respecto a la geometría  $(\mathbb{R}^m, \langle *, * \rangle_P)$  es equivalente al ACP de  $\mathbf{Q}Y$  con respecto a la geometría usual de  $\mathbb{R}^m$ , en el sentido siguiente:*

$$Y(w) = \sum_{j=1}^m \psi_j(w) v_j \quad (\text{con respecto a } (\mathbb{R}^m, \langle *, * \rangle_P))$$

$$\mathbf{Q}Y(w) = \sum_{j=1}^m \psi_j(w) \mathbf{Q}v_j \quad (\text{con respecto a } (\mathbb{R}^m, \langle *, * \rangle_u))$$

Uniendo todos los resultados expuestos anteriormente obtenemos las siguientes consecuencias:

### Corolario 2.3.1

$$1. \mathbf{E}[X] = \gamma(\mathbf{E}[Y])$$

2.  $C_x = \gamma \circ C_y \circ P \circ \gamma^{-1}$

3. *El ACP de la v.a. hilbertiana X con respecto a la geometría del espacio E es equivalente al ACP del v.a. QY con respecto a la geometría usual de R<sup>m</sup>*

*Demostración*

En principio, los dos primeros resultados son consecuencia inmediata de unir lo obtenido en las Proposiciones (2.3.3) y (2.3.5).

Así mismo, el tercer resultado es fruto de unir lo establecido en las Proposiciones (2.3.4) y (2.3.6).

□

**Corolario 2.3.2** *El ACP de la v.a. hilbertiana X dado por*

$$X(w) = \sum_{j=1}^m \psi_j(w) f_j \quad \forall w \in \Omega$$

*se obtiene como sigue:*

- *consideramos la ecuación matricial de diagonalización siguiente*

$$Q \circ C_y \circ Q u_k = \lambda_k u_k \quad \forall k = 1, \dots, m$$

- *asociado al valor propio k-ésimo se obtiene:*

$$\psi_k(w) = u_k^t Q \tilde{Y}(w) \quad \forall w \in \Omega$$

$$f_k = \gamma \circ Q^{-1}(u_k)$$

*donde  $\tilde{Y} \stackrel{\text{def}}{=} Y - E[Y]$ .*

*Demostración*

Obsérvese que el operador de covarianza de QY con respecto a la geometría usual de R<sup>m</sup> puede obtenerse de la forma:

$$C_{QY} = Q \circ C_y \circ Q$$

A partir de dicho operador, el ACP de  $QY$  con respecto a la geometría usual vendría dado por la ecuación matricial siguiente:

$$Q \circ C_y \circ Q u_k = \lambda_k u_k$$

donde estamos denotando con  $u_k$  a un factor (vector) principal de  $QY$  asociado al valor propio  $\lambda_k$ .

En principio, sabemos que  $\lambda_k$  es también un elemento del espectro de  $C_x$ . Además asociado a  $\lambda_k$ , utilizando la Proposición (2.3.6) se obtiene que  $Q^{-1}(u_k)$  sería un factor propio de  $Y$  con respecto a la geometría  $(\mathbb{R}^m, \langle *, * \rangle_P)$ ; de donde a su vez, por la Proposición (2.3.4) se obtendría que  $f_k = \gamma(Q^{-1} u_k)$  es un factor propio de  $X$  con respecto a la geometría del espacio  $\mathbb{E}$ .

Por último, teniendo en cuenta que las componentes principales en todos los ACP anteriormente expuestos son las mismas, se obtiene que la componente principal  $\psi_k$  de la v.a. hilbertiana  $X$  vendría dada por

$$\psi_k(w) = \langle u_k, \widetilde{Q}Y(w) \rangle_u = u_k^t Q \tilde{Y}(w)$$

□

A continuación, mostraremos cómo los resultados aquí obtenidos, para v.a. hilbertianas con valores en un espacio de Hilbert finito dimensional, generalizan algunas de las metodologías desarrolladas en Aguilera (1993) para la puesta en práctica de los métodos de aproximación del ACP de un proceso estocástico. En todas ellas, el común denominador es considerar como aproximación del ACP del proceso, el asociado a un proceso estocástico, que considerado como v.a. hilbertiana tomará valores en un cierto espacio finito dimensional. Dicha v.a. hilbertiana, a la que nos referiremos como *v.a. aproximadora*, dependerá del método numérico que se esté considerando.

### Método de Proyección ortogonal

Aunque un estudio más profundo de este método puede seguirse en Aguilera (1993) pag 82, la metodología que allí se desarrolla se inicia con la elección de un subespacio funcional,  $(\mathbb{E})$ , finito dimensional de  $L_2(T)$  para aproximar las trayectorias, y la elección en él de una base **ortonormal**  $\mathcal{B} = \{e_1, \dots, e_m\}$ . Una vez realizada dicha elección, se *obtiene* la v.a. *aproximadora*, con respecto de los elementos de la base elegida  $\mathcal{B}$ , de la forma siguiente:  $X(w) = \sum_{i=1}^m Y_i(w) e_i \quad \forall w \in \Omega$ , siendo en este caso  $\mathbf{P} = \mathbf{I}_m$ .



En nuestro ambiente general, la Proposición 2.2.2, demostrada en Aguilera (1993), podría verse como una particularización del Teorema (2.3.1) e interpretarse en términos de la existencia de la *equivalencia entre el ACP de X con respecto a  $L_2(T)$ , y el ACP multivariante clásico (con respecto al producto escalar usual) del vector de coordenadas Y*.

**Método de interpolación mediante splines cúbicos naturales**

Brevemente, el esquema de este método, expuesto en detalle en Aguilera (1993) pag 94, se inicia a partir de la consideración de la información disponible acerca del proceso estocástico en cuestión,  $\{x_t(w)\}$ , sobre un cierto conjunto finito de nodos  $\pi_k \stackrel{\text{def}}{=} \{t_0 < \dots < t_k\} \subseteq T$ , descrita por el vector aleatorio

$$\bar{x}(w) = (x''_{t_0}(w), x_{t_0}(w), \dots, x_{t_k}(w), x''_{t_k}(w))^t \quad \forall w \in \Omega$$

Asociado a dicho conjunto de nodos, puede considerarse como subespacio  $\mathbb{E}$  el constituido por los splines cúbicos  $S_3(\pi_k)$ , considerando como geometría la heredada por el producto escalar de  $L_2(T)$ . De dicho espacio vectorial se conoce, además de que su dimensión es  $k + 3$ , una base denominada base de los B-splines, denotada con:  $\{B_{-1}, \dots, B_{k+1}\}$

La idea consiste en considerar como *v.a. aproximadora*, denotada con  $IX$ , a la obtenida al ajustar un spline cúbico a cada trayectoria en base a la información conocida sobre los nodos  $\bar{x}$ , de forma que se obtenga

$$IX(w) = \sum_{i=-1}^{k+1} \gamma_i(w) B_i \quad \forall w \in \Omega$$

donde, para cada  $w \in \Omega$ , el **vector de coordenadas** denotado con

$$\gamma(w) \stackrel{\text{def}}{=} (\gamma_{-1}(w), \dots, \gamma_{k+1}(w))^t$$

viene determinado como solución del sistema regular dado por una matriz  $\mathbf{A} \in M_{k+3}(\mathbb{R})$  de la forma siguiente:

$$\mathbf{A} \gamma(w) = \bar{x}(w) \quad \forall w \in \Omega$$

obteniéndose que  $\gamma(w) = \mathbf{T} \bar{x}(w) \quad \forall w \in \Omega$ , con  $\mathbf{T} \stackrel{\text{def}}{=} \mathbf{A}^{-1}$ .

Asociado al vector aleatorio de observaciones  $\bar{x}$ , y con respecto al producto escalar usual de  $\mathbb{R}^{k+3}$ , se consideran

$$M \stackrel{\text{def}}{=} \mathbf{E}[\bar{x}] \quad \text{y} \quad \Sigma \stackrel{\text{def}}{=} \text{cov}(\bar{x})$$

Consideremos así mismo, las siguientes matrices de  $M_{k+3}(\mathbb{R})$

$$\mathbf{P} \stackrel{\text{def}}{=} \left( \int_T B_i(s) B_j(s) ds : i, j \in \{-1, \dots, k+1\} \right)$$

$$\mathbf{Q} \stackrel{\text{def}}{=} \mathbf{P}^{\frac{1}{2}}$$

así como el isomorfismo lineal  $\Upsilon : \mathbb{R}^m \mapsto S_3(\pi_k)$ ,  $\Upsilon(\alpha) \stackrel{\text{def}}{=} \bar{B}^t \alpha \quad \forall \alpha \in \mathbb{R}^m$ , donde  $\bar{B}(t) = (B_{-1}(t), \dots, B_{k+1}(t))^t \quad \forall t \in T$ .

En primer lugar, de lo visto en la sección anterior se obtienen las equivalencias de las siguientes afirmaciones:

1. *el ACP de  $IX$  con respecto a la geometría de  $L_2(T)$ .*
2. *el ACP de  $\gamma$  con respecto a la geometría de  $(\mathbb{R}^{k+3}, \langle *, * \rangle_P)$ .*
3. *el ACP de  $\mathbf{Q}\gamma$  con respecto a la geometría usual de  $\mathbb{R}^{k+3}$ .*

A su vez, si consideramos la aplicación lineal (filtro)

$$\mathbf{F} : \mathbb{R}^{k+3} \mapsto (\mathbb{R}^{k+3}, \langle *, * \rangle_u)$$

$$\mathbf{F}(v) = \mathbf{Q} \mathbf{T} v \quad \forall v \in \mathbb{R}^{k+3}$$

podemos considerar el producto escalar sobre  $\mathbb{R}^{k+3}$  definido como sigue:

$$\rho(v_1, v_2) \stackrel{\text{def}}{=} \langle \mathbf{F} v_1, \mathbf{F} v_2 \rangle_u = v_1^t \mathbf{T}^t \mathbf{P} \mathbf{T} v_2 \quad \forall v_1, v_2 \in \mathbb{R}^{k+3}$$

Obsérvese que la idea, al considerar esta aplicación, ha sido el intentar que se cumpla  $\mathbf{Q}\gamma = \mathbf{F}\bar{x}$ , con lo que haciendo uso del resultado obtenido en relación al ACP filtrante puede afirmarse entonces que

(\*) *El ACP de  $\mathbf{Q}\gamma = \mathbf{F}\bar{x}$  con respecto a la geometría usual de  $\mathbb{R}^{k+3}$ , es equivalente, al ACP de  $\bar{x}$  con respecto a la geometría del espacio  $(\mathbb{R}^{k+3}, \rho)$ .*

Por otra parte, el operador de covarianza (matriz) de  $\mathbf{F}\bar{x}$ , con respecto a la geometría usual de  $\mathbb{R}^{k+3}$ , vendría dado por:

$$\text{cov}(\mathbf{F}\bar{x}) = \mathbf{F} \Sigma \mathbf{F}^t = \mathbf{Q} \mathbf{T} \Sigma \mathbf{T}^t \mathbf{Q}$$

con lo cuál el ACP de  $IX$ , con respecto a la geometría inducida por  $L_2(T)$ , podría obtenerse de la forma siguiente:

- a partir de la ecuación matricial

$$\mathbf{Q} \mathbf{T} \Sigma \mathbf{T}^t \mathbf{Q} \beta_i = \lambda_i \beta_i \quad \forall i \in \{-1, \dots, k+1\} \text{ con } \beta_i^t \beta_i = 1$$

obtendríamos  $P\sigma(C_{IX}) = P\sigma(C_{F\bar{x}}) = \{\lambda_i : i = -1, \dots, k+1\}$ , siendo además  $\{\beta_i : i = -1, \dots, k+1\}$  un sistema ortonormal de  $\mathbb{R}^{k+3}$  formado por factores principales de  $F\bar{x} = \mathbf{Q}\gamma$ .

- para cada valor propio  $\lambda_i$ , se obtendría que  $\mathbf{Q}^{-1} \beta_i$  resultaría ser el correspondiente factor propio para el ACP del vector aleatorio  $\gamma$  con respecto a  $(\mathbb{R}^k + 3, \langle *, * \rangle_P)$ ; por su parte, el correspondiente factor principal para el ACP de  $IX$  podría obtenerse a través de la expresión:

$$\Upsilon(\mathbf{Q}^{-1} \beta_i) = \bar{\mathbf{B}}^t \mathbf{Q}^{-1} \beta_i$$

- así mismo, para cada valor propio  $\lambda_i$ , la correspondiente componente principal asociada al ACP de  $IX$  podría ser obtenida de la forma siguiente:

$$\langle \beta_i, \mathbf{F}(\bar{x}(w) - M) \rangle_u = \beta_i^t \mathbf{F}(\bar{x}(w) - M) \quad \forall w \in \Omega$$

donde estamos denotando con  $M \stackrel{\text{def}}{=} \mathbf{E}[\bar{x}]$ .

En definitiva, la metodología computacional con la que se concluye en Aguilera (1993) relativa a la puesta en práctica del ACP del spline cúbico interpolador  $IX$ , con respecto a la geometría heredada de  $L_2(T)$ , podría ser a su vez interpretada como la realización del ACP multivariante del vector aleatorio de observaciones  $F\bar{x}$ , con respecto a la geometría de usual de  $\mathbb{R}^{k+3}$ .

### 2.3.1 MODIFICACION GEOMETRICA EN EL ESPACIO DE COORDENADAS

Continuaremos considerando,  $X : (\Omega, \mathcal{A}, P) \mapsto \mathbb{E}$ , una v. a. hilbertiana del espacio  $L_2(\Omega, \mathcal{A}, P; \mathbb{E})$ , siendo  $\mathbb{E}$  un espacio vectorial de dimensión  $m$  sobre el que tenemos definido un producto escalar  $\langle *, * \rangle_E$ .

De esta forma, para una cierta base de dicho espacio  $\mathbb{E}$ ,  $\mathcal{B} = \{e_1, \dots, e_m\}$ , podemos considerar el vector aleatorio de *coordenadas*

$$Y \stackrel{\text{def}}{=} (Y_1, \dots, Y_m)^t \in L_2(\Omega, \mathcal{A}, P; \mathbb{R}^m)$$

con el que la v. a.  $X$  admite la representación lineal siguiente:

$$X(w) = \sum_{i=1}^m Y_i(w) e_i \quad \forall w \in \Omega$$

En estas condiciones, podría plantearse la realización del ACP de  $X$  con respecto a una *nueva* geometría, determinada por un *nuevo* producto escalar  $\rho$  sobre el espacio vectorial  $\mathbb{E}$ .

En principio, por ser el espacio vectorial  $\mathbb{E}$  finito dimensional, y haciendo uso de lo establecido en el Teorema de Tihonov para espacios vectorial topológicos de dimension finita, puede afirmarse que las topologías asociadas a

$\langle *, * \rangle_E$  y  $\rho$ , respectivamente, son coincidentes. Concretamente, podría afirmarse que el producto escalar  $\rho$  sería continuo con respecto a la topología asociada a  $\langle *, * \rangle_E$ . Con ello, nuestro problema sería equivalente a la realización del ACP de  $F \circ X$ , siendo  $F$  un operador definido sobre  $\mathbb{E}$  y con valores en un cierto espacio de Hilbert, obtenido a partir de  $\langle *, * \rangle_E$  y  $\rho$ . Es decir, una forma alternativa para resolver nuestro problema sería la realización del ACP, con respecto a una cierta geometría, de la v.a. siguiente:

$$F \circ X(w) = \sum_{i=1}^m Y_i(w) F(e_i) \quad \forall w \in \Omega$$

Como puede observarse en esta expresión, con la v.a.  $X$  queda transformada su *parte geométrica*, sin embargo, en principio, la *parte probabilística*, dada por el vector aleatorio  $Y$ , permanece igual. Es decir, podría afirmarse que el planteamiento del ACP con respecto a una nueva geometría sobre el espacio  $\mathbb{E}$ , puede ser resuelto a través de una conveniente transformación de la *parte geométrica* de los *datos*. En todo lo anterior y lo que sigue, los términos

*parte probabilística* y *parte probabilística* se utilizan con el único fin de diferenciar; teniéndose siempre presente que lo que aquí se denomina *parte probabilística* tiene evidentemente estructura geométrica.

A diferencia de lo expuesto anteriormente, nos proponemos en esta sección analizar cuál sería el correspondiente ACP inducido en la v.a.  $X$ , (su estructura geométrica) mediante una transformación realizada sobre la *parte probabilística* de la v.a.  $X$ .

En concreto, supóngase que se realiza una transformación del vector aleatorio de coordenadas (por ejemplo se tipifican sus componentes), y a continuación se realiza el ACP de dicha transformación con respecto a la geometría euclídea usual de  $\mathbb{R}^m$ . En estas condiciones, nos planteamos intentar interpretar geoméricamente dicho ACP buscando su equivalencia con algún posible ACP de la v.a.  $X$  original.

De la sección anterior sabemos que: *el ACP de  $X$  en el espacio  $(\mathbb{E}, \langle *, * \rangle_E)$  es equivalente al ACP multivariante de  $Y$  en el espacio  $(\mathbb{R}^m, \langle *, * \rangle_P)$* , siendo

$$P = (\langle e_i, e_j \rangle_E : i, j = 1, \dots, m) \in M_m(\mathbb{R}^m)$$

y notando con  $Q = P^{\frac{1}{2}}$ .

Por su parte, dada una matriz  $D \in M_m(\mathbb{R})$  no singular, con el vector aleatorio  $DY$  podría llevarse a cabo el ACP con respecto a la geometría usual de  $\mathbb{R}^m$ . A continuación, estudiaremos la posibilidad de interpretar dicho análisis en términos de un ACP de  $X$  con respecto a una cierta geometría, a obtener, sobre el espacio vectorial  $\mathbb{E}$ .

**Corolario 2.3.3** *En las condiciones anteriores, son equivalentes las siguientes afirmaciones:*

1. *El ACP de  $DY$  con respecto a  $(\mathbb{R}^m, \langle *, * \rangle_u)$ .*
2. *El ACP de  $Y$  con respecto a  $(\mathbb{R}^m, \langle *, * \rangle_K)$   
donde  $K = D^t D = (k_{ij}) \in M_m(\mathbb{R})$*
3. *El ACP de  $X$  con respecto a  $(\mathbb{E}, \rho)$*

*siendo  $\rho$  un producto escalar sobre  $\mathbb{E}$  definido por:*

$$\rho(e_i, e_j) = k_{ij} \quad \forall i, j \in \{1, \dots, m\}$$

□

Por lo obtenido en el Corolario anterior, el *nuevo* producto escalar  $\rho$  estaría definido como sigue:

$$\rho(f, g) = \bar{f}^t \mathbf{K} \bar{g} \quad \forall f, g \in \mathbb{E}$$

siendo  $\bar{f} \stackrel{\text{def}}{=} (f_1, \dots, f_m)^t$  y  $\bar{g} \stackrel{\text{def}}{=} (g_1, \dots, g_m)^t$  vectores de  $\mathbb{R}^m$  tales que

$$f = e^t \bar{f} \quad \text{y} \quad g = e^t \bar{g}$$

donde notamos con  $e \stackrel{\text{def}}{=} (e_1, \dots, e_m)^t$ , permitiendo así utilizar la reglas de cálculo matricial en las expresiones que irán apareciendo con un ahorro en la notación.

Por otro lado, por ser el nuevo producto escalar  $\rho$  continuo respecto a la topología del espacio  $(\mathbb{E}, \langle *, * \rangle_E)$ , puede garantizarse la existencia de un operador  $T : (\mathbb{E}, \langle , \rangle_E) \mapsto (\mathbb{E}, \langle , \rangle_E)$  cumpliendo que:

$$\langle e_i, T(e_j) \rangle_E = \rho(e_i, e_j) = k_{ij} \quad \forall i, j \in \{1, \dots, m\}$$

De esta forma, denotando con  $\alpha_j \stackrel{\text{def}}{=} (\alpha_{1j}, \dots, \alpha_{mj})^t$ , puede considerarse, para cada  $j \in \{1, \dots, m\}$ , al elemento  $T(e_j)$  expresado como sigue:

$$T(e_j) = \sum_{l=1}^m \alpha_{lj} e_l = e^t \alpha_j$$

Con lo cuál, dichos vectores  $\{\alpha_j\}$  quedarían caracterizados por el hecho de verificarse:

$$k_{ij} = \langle e_i, T(e_j) \rangle_E = \sum_{l=1}^m \langle e_i, e_l \rangle_E \alpha_{lj} \quad \forall i, j \in \{1, \dots, m\}$$

o equivalentemente en forma matricial dicha expresión quedaría:

$$\mathbf{K} = \mathbf{P} \mathbf{A} \implies \mathbf{A} = \mathbf{P}^{-1} \mathbf{K}$$

donde estamos notando con  $\mathbf{A} = [\alpha_1 | \dots | \alpha_m] \in M_m(\mathbb{R})$ .

Por tanto, dado  $f = e^t \bar{f}$ , se obtiene que

$$T(f) = \sum_{j=1}^m f_j T(e_j) = \sum_{j=1}^m f_j e^t \alpha_j = e^t \sum_{j=1}^m \alpha_j f_j = e^t \mathbf{A} \bar{f}$$

Así mismo, si consideramos la aplicación  $L : (\mathbb{E}, \langle \cdot, \cdot \rangle_E) \mapsto (\mathbb{E}, \langle \cdot, \cdot \rangle_E)$  definida con,  $L(f) = e^t \mathbf{B} \bar{f} \quad \forall f \in \mathbb{E}$ , donde con  $\mathbf{B} \stackrel{\text{def}}{=} \mathbf{Q}^{-1} \mathbf{D}$ , puede comprobarse que se cumple:

$$\begin{aligned} \langle L(f), L(g) \rangle_E &= \bar{f}^t \mathbf{B}^t \langle e e^t \rangle_E \mathbf{B} \bar{g} = \bar{f}^t \mathbf{D}^t \mathbf{Q}^{-1} \mathbf{P} \mathbf{Q}^{-1} \mathbf{D} \bar{g} = \\ &= \bar{f}^t \mathbf{K} \bar{g} = \langle T(f), g \rangle_E \end{aligned}$$

Concluyéndose así que  $T = L^* \circ L$ , siendo  $L^*$  el operador adjunto del operador  $L$  del espacio  $(\mathbb{E}, \langle \cdot, \cdot \rangle_E)$  en sí mismo.

**Proposición 2.3.7** *Teniendo en cuenta lo anterior, las siguientes afirmaciones son equivalentes:*

1. El ACP de  $\mathbf{D}Y$  con respecto a  $(\mathbb{R}^m, \langle \cdot, \cdot \rangle_u)$ .
2. El ACP de  $X$  con respecto a  $(\mathbb{E}, \rho)$ .
3. El ACP de  $L \circ X$  con respecto a  $(\mathbb{E}, \langle \cdot, \cdot \rangle_E)$   
donde  $L(f) = e^t \mathbf{Q}^{-1} \mathbf{D} \bar{f} \quad \forall f = e^t \bar{f} \in \mathbb{E}$

Obsérvese que las componentes  $\{v_j\}$  que constituyen el vector  $v \in \mathbb{E}^m$  dado por

$$v^t = (v_1, \dots, v_m) \stackrel{\text{def}}{=} e^t \mathbf{Q}^{-1}$$

forman una base ortonormal del espacio  $(\mathbb{E}, \langle \cdot, \cdot \rangle_E)$ , ya que como puede comprobarse se tiene:

$$\langle v v^t \rangle_E = \mathbf{Q}^{-1} \langle e e^t \rangle_E \mathbf{Q}^{-1} = \mathbf{Q}^{-1} \mathbf{P} \mathbf{Q}^{-1} = I_m$$

es decir  $\langle v_i, v_j \rangle_E = \delta_{ij} \quad \forall i, j \in \{1, \dots, m\}$ .

De esta forma, dado  $f = e^t \bar{f} \in \mathbb{E}$ , obsérvese que se tendría

$$f = e^t \bar{f} = e^t \mathbf{Q}^{-1} \mathbf{Q} \bar{f} = v^t (\mathbf{Q} \bar{f})$$

con lo que el vector de coordenadas de  $f$  respecto a la base, formada por las componentes de  $v$ , vendría dado por  $\mathbf{Q} \bar{f}$ .

En estas condiciones, el operador  $L$  podría interpretarse de la forma siguiente:

$$\begin{aligned} L : \mathbb{E} &\longrightarrow \mathbb{E} \\ f = e^t \bar{f} &\longmapsto v^t \mathbf{D} \bar{f} \end{aligned}$$

**Ejemplo 2.3.1** En el caso de que la base  $\mathcal{B} = \{e_1, \dots, e_m\}$  sea ortonormal en el espacio  $(\mathbb{E}, \langle *, * \rangle_E)$ , el operador  $L$  queda de la forma siguiente:

$$\begin{aligned} L : \mathbb{E} &\longrightarrow \mathbb{E} \\ f = e^t \bar{f} &\longmapsto e^t \mathbf{D} \bar{f} \end{aligned}$$

**Ejemplo 2.3.2** Supóngase que la base  $\mathcal{B}$  es ortonormal en el espacio  $(\mathbb{E}, \langle *, * \rangle_E)$ , y consideremos el caso de la tipificación es decir

$$\mathbf{D} = \text{diag}(\sigma_1^{-1}, \dots, \sigma_m^{-1}) \in M_m(\mathbb{R})$$

entonces

$$L \circ X(w) = \sum_{j=1}^m \sigma_j^{-1} Y_j(w) e_j \quad \forall w \in \Omega$$



## Capítulo 3

# MODELIZACION MEDIANTE ACP

### §3.1 PREDICCIÓN LINEAL EN COMPONENTES PRINCIPALES

#### 3.1.1 NOTACION Y TEORIA BASICA

Acerca del problema de la predicción, de un proceso estocástico en tiempo continuo en términos de su pasado reciente, puede encontrarse un considerable número de enfoques a todo largo de la literatura estadística y probabilística. Muchos de estos estudios proporcionan una solución mediante la imposición de, entre otras restricciones, la estacionaridad. En esta línea de estudio, Bosq (1991) introdujo los procesos hilbertianos autoregresivos (ARH) como una generalización de los procesos clásicos AR a v.a. con valores en un espacio funcional de Hilbert.

Teniendo en cuenta que un gran número de aplicaciones no cumplen las hipótesis de dichos modelos particulares, se aplicará el ACP con objeto de reducir la dimensión de los espacios de las variables de respuesta y predictor, con la consiguiente simplificación en la solución. Por otra parte, para el caso particular de espacios de Hilbert con núcleos reproductores, y basándose en los trabajos de Besse y Ramsay (1986) y Besse (1991), Besse y Cardot (1994) han desarrollado un modelo de predicción para v.a. hilbertianas ARH(1).

La técnica de predicción que se va a proponer está basada en la regresión lineal de las componentes principales (cc.pp.) asociadas al proceso en el

futuro con respecto a las correspondientes componentes asociadas al pasado. Esta puede verse como una extensión de la técnica de Regresión Multivariante en Componentes Principales de múltiples respuestas (RMCP, see e.g., Jolliffe (1985)) para predecir un conjunto infinito de respuestas a partir de un conjunto infinito de predictores.

### REGRESION MULTIVARIANTE EN COMPONENTES PRINCIPALES

La técnica RMCP resuelve el problema de predicción de un conjunto de variables de respuestas, dadas en el vector aleatorio  $\underline{Y} = (Y_1, \dots, Y_s)'$ , a partir de predictores, denotados con  $\underline{X} = (X_1, \dots, X_m)'$  ( $m, s \in \mathbb{Z}^+$ ).

La principal característica de esta técnica es que se realiza un doble ACP. Es decir, se realiza un ACP clásico para cada uno de los conjuntos de variables anteriormente citados.

Denotemos con  $\{\xi_1, \dots, \xi_m\}$  las componentes principales asociadas al vector aleatorio  $\underline{X}$ , dado por

$$\xi_i = \sum_{j=1}^m \phi_{ji} X_j = \underline{\phi}_i' \underline{X}, \quad (3.1)$$

donde  $\underline{\phi}_i$  es el  $i$ -ésimo vector propio de la matriz de covarianza de  $\underline{X}$ .

De la misma forma, denotemos con  $\{\eta_1, \dots, \eta_s\}$  las cc.pp. asociadas al vector aleatorio  $\underline{Y}$ , definidas como sigue

$$\eta_k = \sum_{j=1}^s \delta_{jk} Y_j = \underline{\delta}_k' \underline{Y}, \quad (3.2)$$

donde  $\underline{\delta}_k$  es el  $k$ -ésimo vector propio de la matriz de covarianza de  $\underline{Y}$ .

Entonces, un modelo RMCP para predecir cada variable de respuesta  $Y_j$  ( $j = 1, \dots, s$ ) viene dado por

$$\tilde{Y}_j = \bar{Y}_j + \sum_{k=1}^q \hat{\eta}_k \delta_{jk}, \quad q \leq s, \quad (3.3)$$

donde  $\hat{\eta}_k$  ( $k = 1, \dots, s$ ) es el estimador lineal mínimo cuadrático de la c.p.  $\eta_k$  sobre un subconjunto del conjunto de c.c.p.p.  $\{\xi_i\}_{i=1}^m$ .

Si en la ecuación (3.3)  $q$  es igual a  $s$ , todas las c.p.'s  $\{\xi_i\}_{i=1}^m$  son usadas para estimar  $\eta_k$ , con lo que RMCP producirá los mismos resultados que en

la regresión lineal mínimo cuadrática multivariantes (RLMM), salvo con la posibilidad de una mayor exactitud en casos en los que existan problemas de inversión para la matriz de covarianzas muestrales.

### REGRESION FUNCIONAL EN COMPONENTES PRINCIPALES

Denotemos con  $[T_1, T_2]$  y  $[T_3, T_4]$  ( $T_3 > T_2$ ) a los intervalos de tiempo, pasado y futuro respectivamente. Consideremos ahora un proceso aleatorio de segundo orden continuo en media cuadrática,  $\{X(t)\}$ , cuyas funciones muestrales son de cuadrado integrable de Lebesgue sobre dichos intervalos  $[T_1, T_2]$  y  $[T_3, T_4]$ .

El problema consiste en encontrar un estimador lineal de la variable aleatoria  $X(s)$  ( $s \in [T_3, T_4]$ ) a partir de las variables  $\{X(t) : t \in [T_1, T_2]\}$ . Como el estimador propuesto será función de las componentes principales de  $\{X(t)\}$ , comenzaremos considerando la teoría básica acerca del ACP de un proceso aleatorio, respecto a la geometría usual de  $L_2[T_1, T_2]$ .

En principio, la  $i$ -ésima componente principal asociada al proceso  $\{X(t)\}$  en el intervalo  $[T_1, T_2]$  viene definida por:

$$\xi_i = \int_{T_1}^{T_2} (X(t) - \mu(t)) f_i(t) dt, \tag{3.4}$$

donde  $f_i$ , denominado  $i$ -ésimo factor principal (f.p.), es la función propia normalizada correspondiente al  $i$ -ésimo mayor valor propio  $\lambda_i$  del operador de covarianza con núcleo  $C(t, s)$ , y denotando con  $\mu(t)$  la función media del proceso.

Con lo cuál, la varianza explicada por la  $i$ -ésima componente principal es  $V_i^P = \lambda_i/V^P$ , siendo  $V^P = \sum_i \lambda_i$  la varianza total del proceso en el intervalo pasado  $[T_1, T_2]$ .

En estas condiciones, puede considerarse la siguiente descomposición en cc.pp. para el proceso en el intervalo pasado:

$$X(t) = \mu(t) + \sum_{i=1}^{\infty} \xi_i f_i(t) \quad t \in [T_1, T_2], \tag{3.5}$$

donde la serie que aparece en (3.5) converge en media cuadrática a  $X(t)$ , uniformemente en  $t$ .

Además, la serie (3.5) truncada en el  $q$ -ésimo término es el mejor estimador lineal mínimo cuadrático  $q$ -dimensional para el proceso centrado  $X(t) - \mu(t)$

(ver, por ejemplo, Fukunaga (1990)), siendo  $(\sum_{i=1}^q \lambda_i)/V$  la proporción de varianza explicada por este modelo y  $\sum_{i=q+1}^{\infty} \lambda_i$  el error mínimo cuadrático.

De forma similar, puede considerarse la correspondiente representación ortogonal asociada al ACP del proceso en el intervalo futuro:

$$X(s) = \mu(s) + \sum_j \eta_j g_j(s) \quad s \in [T_3, T_4] \quad (3.6)$$

donde con  $g_j$  y  $\eta_j$  denotamos los factores y componentes principales en el intervalo  $[T_3, T_4]$ , respectivamente.

Finalmente, la varianza total del proceso en el futuro viene dada por  $V^F = \sum_j \alpha_j$ , siendo  $\alpha_j$  la varianza de la c.p.  $\eta_j$ . Con lo cuál, la proporción de varianza explicada por la  $j$ -ésima c.p.  $\eta_j$  viene dada por  $V_j^F = \alpha_j/V^F$ .

### 3.1.2 MODELO DE PREDICCIÓN EN COMPONENTES PRINCIPALES

Consideremos  $\{X(t)\}$  un proceso estocástico centrado y definido sobre un espacio probabilístico  $(\Omega, \mathcal{A}, P)$ . Sea  $L_X^2$  el subespacio lineal cerrado generado por las vv.aa.  $\{X(t) : T_1 \leq t \leq T_2\}$ , es decir

$$L_X^2 = \overline{\text{Lin}\{X(t) : t \in [T_1, T_2]\}}^{L_2(\Omega, \mathcal{A}, P)}.$$

En principio, estamos interesados en la predicción lineal de  $X(s)$  ( $s \in [T_3, T_4]$ ), dados los valores del proceso en el periodo  $[T_1, T_2]$ . Por tanto, el mejor predictor lineal de  $X(s)$  queda definido como aquella variable aleatoria  $\tilde{X}(s)$  verificando

$$E[|\tilde{X}(s) - X(s)|^2] = \inf\{E[|Z - X(s)|^2] : Z \in L_X^2\}.$$

En la terminología de los espacios de Hilbert,  $\tilde{X}(s)$  es la proyección ortogonal de  $X(s)$  en  $L_X^2$ , caracterizada como el elemento de  $L_X^2$  tal que

$$\tilde{X}(s) - X(s) \perp L_X^2,$$

o equivalentemente

$$E[X(s)X(t)] = E[\tilde{X}(s)X(t)] \quad \forall t \in [T_1, T_2],$$

A la cantidad

$$\epsilon^2(s) \stackrel{\text{def}}{=} E[|\tilde{X}(s) - X(s)|^2]$$

se le denomina error en media cuadrática de la predicción lineal en  $s \in [T_3, T_4]$ .

Por otro lado, se sabe (ver Todorovic (1992)) que las c.c.p.p.  $\{\xi_i\}$ , asociadas al proceso en el pasado, constituyen una base ortogonal del  $L^2_X$ , verificando:  $E[\xi_i^2] = \lambda_i$ . Con lo que la proyección ortogonal  $P : L^2(\Omega) \rightarrow L^2_X$  admitiría la siguiente expresión, convergente en media cuadrática, en termino de dichas cc.pp. normalizadas del pasado:

$$P(\phi) = \sum_{i=1}^{\infty} \frac{E[\phi \xi_i]}{\lambda_i} \xi_i \quad \forall \phi \in L_2(\Omega, \mathcal{A}, P). \quad (3.7)$$

Debido al hecho de que  $P$  es un operador lineal continuo sobre  $L_2(\Omega, \mathcal{A}, P)$  es fácil probar, a partir de las ecuaciones (3.6) y (3.7), que el estimador lineal mínimo cuadrático  $\tilde{X}(s)$  puede ser expresado como

$$\tilde{X}(s) = P(X(s)) = \sum_{j=1}^{\infty} \tilde{\eta}_j g_j(s), \quad s \in [T_3, T_4], \quad (3.8)$$

donde  $\tilde{\eta}_j$  viene dado por

$$\tilde{\eta}_j \stackrel{\text{def}}{=} P(\eta_j) = \sum_{i=1}^{\infty} \frac{E[\eta_j \xi_i]}{\lambda_i} \xi_i = \sum_{i=1}^{\infty} \beta_i^j \xi_i \quad j = 1, \dots, \quad (3.9)$$

Obsérvese que  $\tilde{\eta}_j$  representa al estimador lineal mínimo cuadrático de la c.p.  $\eta_j$  con respecto a las variables del proceso  $\{X(t) : T_1 \leq t \leq T_2\}$ .

Mediante un truncamiento en la serie de la ecuación (3.9), se obtiene la siguiente predicción lineal *aproximada* para cada una de las cc.pp. en el futuro

$$\tilde{\eta}_j^{p_j} = \sum_{i=1}^{p_j} \beta_i^j \xi_i. \quad (3.10)$$

Finalmente, podemos construir el siguiente modelo predictivo en componentes principales para el proceso en el futuro

$$\tilde{X}^q(s) = \mu(s) + \sum_{j=1}^q \tilde{\eta}_j^{p_j} g_j(s) \quad s \in [T_3, T_4], \quad (3.11)$$

donde la ecuación (3.6) ha sido restringida para las  $q$  cc.pp. en el futuro cuya varianza acumulada  $CV_j^F = \sum_{j=1}^q \alpha_j$ . A este modelo lo denotaremos con  $PCP(q, p_1, \dots, p_q)$ .

Por último, el error mínimo cuadrático total para el modelo PCP (3.11) en el intervalo  $[T_3, T_4]$  estaría definido como

$$\epsilon^2 = E \left[ \int_{T_3}^{T_4} (X(s) - \tilde{X}^q(s))^2 ds \right]. \quad (3.12)$$

Después de algunos calculos puede obtenerse la siguiente expresión para  $\epsilon^2$ , en terminos de los valores principales y los coeficientes de regresion, concretamente:

$$\begin{aligned} \epsilon^2 &= \sum_{j=1}^{\infty} \alpha_j - \sum_{j=1}^q \sum_{i=1}^{p_j} \frac{1}{\lambda_i} (E[\eta_j, \xi_i])^2 \\ &= V^F - \sum_{j=1}^q \sum_{i=1}^{p_j} \lambda_i (\beta_i^j)^2 \end{aligned} \quad (3.13)$$

Resumiendo, de las ecuaciones (3.10) y (3.11) claramente se observa que nuestro problema ha quedado reducido a realizar la regresión lineal mínimo cuadrática para algunas de las primeras cc.pp. en el futuro sobre algunas de las primeras cc.pp. del pasado.

En las aplicaciones, el problema principal es elegir el número óptimo  $p_j$  de cc.pp.  $\{\xi_i\}$  para ser introducidas en el modelo PCP como predictoras para cada una de las primeras  $q$  cc.pp.  $\eta_j$  en el futuro. Aunque la práctica usual consiste en seleccionar automáticamente como predictoras, aquellas cc.pp. asociadas a los mayores valores propios, algunos autores como Hötelling y Jolliffe, entre otros, (ver, por ejemplo, Jackson (1992)), señalan, en el caso multivariante, que no existe ninguna razón para que sean los mejores predictores las cc.pp. con máxima varianza, dando ejemplos donde alguna de las cc.pp. con menor varianza presentaban una alta correlación con la variable de respuesta. Por esta razón, elegiremos aquellas cc.pp. en el pasado que presenten una correlación significativa para cada una de las cc.pp. del futuro.

### 3.1.3 AJUSTE DE LOS MODELOS DE PREDICCIÓN

Una vez definido el modelo PCP, procederemos a estimarlo a partir de  $N$  funciones muestrales independientes del proceso, en cada uno de los intervalos, denotadas con

$$\{X_w(t) : t \in [T_1, T_2]\} \quad \text{y} \quad \{X_w(s) : s \in [T_3, T_4]\} \quad (w = 1, \dots, N).$$

En estas condiciones, la identificación, estimación y diagnóstico de los modelos PCP se realiza siguiendo los siguientes pasos:

1. Tendremos que estimar los factores principales del proceso en cada intervalo. Es sabido (Deville (1973)) que los estimadores naturales de los f.p.'s, a partir de un conjunto de trayectorias muestrales independientes, son las funciones propias del operador covarianza muestral con núcleo dado por

$$\hat{C}(t, s) = \frac{1}{N-1} \sum_{w=1}^N (X_w(t) - \bar{X}(t)) (X_w(s) - \bar{X}(s)), \quad (3.14)$$

siendo  $\bar{X}$  el estimador usual insesgado de la media  $\mu$ .

Los valores y factores principales muestrales asociados al proceso  $\{X(t)\}$  en el pasado, denotados con  $\hat{\lambda}_i$  y  $\hat{f}_i$  respectivamente, son las soluciones de la ecuación integral de Fredholm de segunda especie:

$$\int_{T_1}^{T_2} \hat{C}(t, s) \hat{f}_i(s) ds = \hat{\lambda}_i \hat{f}_i(t) \quad \forall t \in [T_1, T_2]. \quad (3.15)$$

Una vez que los factores principales han sido estimados, la  $i$ -ésima c.p. muestral asociada al proceso en el pasado  $[T_1, T_2]$  vendría dada por

$$\hat{\xi}_{wi} = \int_{T_1}^{T_2} (X_w(t) - \bar{X}(t)) \hat{f}_i(t) dt \quad w = 1, \dots, N. \quad (3.16)$$

De igual forma, con  $\hat{\alpha}_j$ ,  $\hat{g}_j$  y  $\hat{\eta}_j$  denotaremos a los valores, factores y componentes principales muestrales en el intervalo futuro  $[T_3, T_4]$ , respectivamente.

2. Para elegir el número optimal  $q$  de cc.pp.  $\hat{\eta}_j$  para ser introducidas en el modelo PCP como variables de respuesta, fijaremos una proporción (aproximadamente entre 80 y 99 %) y consideraremos las primeras  $q$  cc.pp. cuyo porcentaje de varianza acumulativa sea mayor o igual a esta proporción.
3. Para seleccionar los mejores predictores para una c.p. en el futuro, mediremos las correlaciones lineales muestrales entre cada c.p. en el pasado



y dicha c.p. en el futuro. Por tanto, para cada c.p.  $\hat{\eta}_j$  ( $j = 1, \dots, q$ ) introduciremos en el modelo de regresión (3.10) las cc.pp.  $\{\hat{\xi}_i\}$  en orden a la magnitud del cuadrado de sus correlaciones con la variable de respuesta  $\hat{\eta}_j$ .

4. Una vez que las cc.pp. de respuesta y predictoras han sido identificadas, estimaremos, en la forma usual, el modelo de regresión para cada c.p.  $\hat{\eta}_j$  ( $j = 1, \dots, q$ ) con respecto a sus predictoras  $\hat{\xi}_i$  ( $i = 1, \dots, p_j$ )

$$\tilde{\eta}_j^{p_j} = \sum_{i=1}^{p_j} \hat{\beta}_i^j \hat{\xi}_i$$

donde el vector de coeficientes de regresión  $\hat{\beta}^j$  viene dado por

$$\hat{\beta}^j = [\Xi' \Xi]^{-1} \Xi' \hat{\eta}_j \quad (3.17)$$

y

- $\Xi = (\hat{\xi}_1, \dots, \hat{\xi}_{p_j})$  es la matriz de regresión de dimensión  $N \times p_j$  cuyas columnas son los vectores

$$\hat{\xi}_i = (\hat{\xi}_{1i}, \dots, \hat{\xi}_{Ni})' \quad (i = 1, \dots, p_j).$$

- $\hat{\eta}_j$  es el vector columna  $N \times 1$  cuyas componentes son los valores muestrales de la c.p.  $\hat{\eta}_j$ . Es decir,

$$\hat{\eta}_j = (\hat{\eta}_{1j}, \dots, \hat{\eta}_{Nj})'$$

Obsérvese que la matriz  $\Xi' \Xi$  es diagonal, con lo que su inversa vendría dada a partir de los inversos de los elementos de la diagonal.

5. Estimaremos el modelo PCP identificado como

$$\tilde{X}^q(s) = \bar{X}(s) + \sum_{j=1}^q \tilde{\eta}_j^{p_j} \hat{g}_j(s) \quad s \in [T_3, T_4]. \quad (3.18)$$

Para cada nueva observación individual  $w$  observada exclusivamente en el intervalo pasado  $[T_1, T_2]$ , obtenemos una predicción de  $\tilde{X}_w^q(s)$  para cualquier  $s \in [T_3, T_4]$ , una vez estimadas sus componentes principales en el pasado y rempazandolas en la ecuación (3.18).



6. Por último, para evaluar lo adecuado del modelo PCP estimado estimaremos el error cuadrático medio como

$$\hat{\epsilon}^2(s) = \frac{1}{N-1} \sum_{w=1}^N (X_w(s) - \tilde{X}^q(s))^2 \quad s \in [T_3, T_4], \quad (3.19)$$

y el error cuadrático medio total como

$$\hat{\epsilon}^2 = \hat{V}^F - \sum_{j=1}^q \sum_{i=1}^{p_j} \hat{\lambda}_i (\hat{\beta}_i^j)^2 \quad (3.20)$$

con  $\hat{V}^F = \sum_j \hat{\alpha}_j$

### §3.2 MODELIZACION CON DATOS DISCRETOS

Desde Deville (1974), que estableció la teoría básica del ACP de procesos continuos, un gran número de estudios se han encaminado a tratar el problema de la aproximación del ACP a partir de un conjunto de trayectorias independientes, y en especial al caso en que la información disponible de dichas curvas venga tan sólo dada por sus valores en un conjunto discreto de instantes. En tales circunstancias, una aproximación alternativa podría ser la realización del ACP clásico de la matriz de datos observados. En Castro et al. (1986) se analiza el efecto sobre los resultados del ACP cuando las trayectorias muestrales son sólo conocidas sobre nodos desigualmente espaciados. Ellos usaron fórmulas de cuadratura para aproximar el ACP dieron ejemplos que mostraban que el buen comportamiento de robustez de los resultados proporcionados por dicha aproximación funcional, frente a los obtenidos utilizando el ACP clásico. En ambientes de espacios de Hilbert con núcleo reproductor, Besse y Ramsay (1986) y Besse (1991), utilizan *splines interpoladores* y *aproximadores* para llevar a cabo dicho análisis.

En esta sección, se consideraran dos técnicas numéricas para la aproximación del ACP de un proceso: aproximación de las trayectorias muestrales mediante proyección ortogonal en un subespacio de dimensión finita (Aguilera et al. (1995)), y la interpolación spline cúbica sobre los instantes de observación (Aguilera (1993)).

En muchos contextos aplicados dispondremos sólo de un conjunto de observaciones discretas, que denotaremos con  $\{X_w(t_j) : w = 1, \dots, N; j = 0, 1, \dots, m\}$ ,

y que representarán toda la información disponible relativa a la correspondiente muestra,  $\{X_w(t) : w = 1, \dots, N\}$ , de trayectorias independientes del proceso  $\{X(t)\}$ .

Por ejemplo, para el anterior problema de predicción del proceso en el instante  $s \in [t_{k+1}, t_m]$ , ( $k : 1, \dots, m - 2$ ), consideraremos en el periodo de observación  $[t_0, t_m]$  dos intervalos disjuntos,  $[t_0, t_k]$  y  $[t_{k+1}, t_m]$ , representando el pasado y el futuro, respectivamente, y considerando  $T_1 = t_0$ ,  $T_2 = t_k$ ,  $T_3 = t_{k+1}$  and  $T_4 = t_m$ .

En estas condiciones, los factores principales muestrales han de ser aproximados en cada intervalo utilizando técnicas numericas eficientes. Por ejemplo. Besse y Ramsay (1986) resuelven este problema suponiendo como hipótesis que las funciones muestrales están contenidas en un determinado espacio de Hilbert con núcleo reproductor.

## PROYECCION ORTOGONAL

El método de proyección ortogonal (MPO) aproxima los factores principales a través de los correspondientes factores del proceso estocástico obtenido al proyectar las trayectorias del proceso original sobre un subespacio funcional de dimensión finita  $E_n$  de  $L^2[T_1, T_2]$ .

En Aguilera et al. (1993) se demuestra que para una base ortonormal de  $E_n$ ,  $\{e_j\}_{j=1}^n$ , el  $i$ -ésimo factor principal aproximado, denotado con  $\hat{f}_i^{(n)}$ , viene dado por

$$\hat{f}_i^{(n)} = \sum_{j=1}^n \gamma_{ji} e_j \quad (3.1)$$

siendo el vector  $\underline{\gamma}_i$ , con componentes  $\gamma_{ji}$ , solución de la ecuación matricial

$$\mathbf{R} \underline{\gamma}_i = \hat{\lambda}_i^{(n)} \underline{\gamma}_i,$$

donde con  $\mathbf{R}$  estamos denotando a la matriz  $n \times n$  con elementos

$$R_{ij} = \int_{T_1}^{T_2} \int_{T_1}^{T_2} \hat{C}(t, s) e_i(t) e_j(s) dt ds = \frac{1}{N-1} \sum_{w=1}^N (Y_{wi} - \bar{Y}_i)(Y_{wj} - \bar{Y}_j)$$

definiendo a su vez

$$Y_{wj} = \int_{T_1}^{T_2} X_w(s) e_j(s) ds \quad \forall j = 1, \dots, n, \quad (3.2)$$

$$\bar{Y}_j = \frac{1}{N} \sum_{w=1}^N Y_{wj} = \int_{T_1}^{T_2} \bar{X}(s) e_j(s) ds.$$

Con ello, las componentes principales muestrales aproximadas vendrán dadas por

$$\hat{\xi}_{wi}^{(n)} = \sum_{j=1}^n \gamma_{ji} (Y_{wj} - \bar{Y}_j) \quad w = 1, \dots, N. \quad (3.3)$$

Obsérvese como estas cc.pp. muestrales aproximadas resultan ser también las de la matriz de integrales  $\mathbf{Y}$ , cuyos elementos vienen definidos en la ecuación (3.2).

En relación a la elección del subespacio  $E_n$ , señalar brevemente que esta se lleva a cabo teniendo en cuenta la naturaleza de las trayectorias del proceso. Por ejemplo, para procesos con trayectorias constantes en intervalos aleatorios, un subespacio  $E_n$  a considerar podría ser el formado por las funciones constantes en intervalos determinados por una partición previamente fijada en  $[T_1, T_2]$  (Deville (1974)). Por su parte, en el caso de trayectorias muestrales regulares podría considerarse el subespacio engendrado por las funciones trigonométricas (Aguilera et al. (1993)). En el artículo de Aguilera et al. (1995), los autores analizan la exactitud de este método a través de la comparación de resultados obtenidos a partir de observaciones discretas simuladas para procesos cuyo ACP es conocido de forma exacta.

### INTERPOLACION SPLINE CUBICA

En el caso de trayectorias muestrales regulares, un método alternativo es *aproxi-  
mar* los ff.pp. muestrales a través de los ff.pp. correspondientes al conjunto de splines cúbicos naturales ajustados a las observaciones discretas disponibles para cada una de las trayectorias muestrales. Nos referiremos a este método con IBC.

Denotando con  $\{B_p\}_{p=-1}^{k+1}$  a los B-splines cúbicos sobre  $[T_1, T_2]$ , con nodos  $t_j$  ( $j = 0, \dots, k$ ). En Aguilera (1993) se prueba que el  $i$ -ésimo f.p. del conjunto de trayectorias muestrales interpoladas, denotado con  $\hat{f}_i^I$ , es un spline cúbico natural dado por

$$\hat{f}_i^I(t) = \sum_{p=-1}^{k+1} \alpha_{pi} B_p(t) \quad (3.4)$$

donde el vector columna  $\underline{\alpha}_i$ , cuyos elementos son  $\alpha_{pi}$  ( $p = -1, \dots, k+1$ ), se obtiene a partir de la ecuación de diagonalización

$$\mathbf{T}\hat{\Sigma}\mathbf{T}'\mathbf{P}\underline{\alpha}_i = \hat{\lambda}_i^I \underline{\alpha}_i, \quad (3.5)$$

siendo  $\mathbf{P}$  la matriz  $(k+3) \times (k+3)$  cuyos elementos son los productos escalares entre los B-splines,  $\mathbf{T}$  es la inversa de la matriz de interpolación, y  $\hat{\Sigma}$  es la matriz de covarianzas muestrales del vector aleatorio  $(X''(t_0), X(t_0), \dots, X(t_k), X''(t_k))'$ .

Además, definiendo  $\underline{\alpha}_i = \mathbf{P}^{-1/2}\underline{\beta}_i$ , el problema se reduce a calcular los vectores propios  $\underline{\beta}_i$  de la matriz simétrica  $\mathbf{P}^{1/2}\mathbf{T}\hat{\Sigma}\mathbf{T}'\mathbf{P}^{1/2}$ .

Finalmente, las cc.pp. pueden obtenerse a partir de la siguiente expresión

$$\hat{\xi}_{wi}^I = \sum_{p=-1}^{k+1} (a_{wp} - \bar{a}_p) \sum_{s=-1}^{k+1} P_{ps} \alpha_{si}, \quad (3.6)$$

donde los  $a_{wp}$  son, para cada  $w = 1, \dots, N$ , los coeficientes de interpolación para la trayectoria muestral  $X_w(t)$  en la base de los B-splines basis; y denotando con

$$\bar{a}_p = \frac{1}{N} \sum_{w=1}^N a_{wp}.$$

Por último, resaltar que en la práctica ocurre normalmente que no pueden observarse las derivadas segundas,  $X''(t_0)$  y  $X''(t_k)$ , para las curvas muestrales. Por esta razón, en la aproximación del ACP del proceso mediante interpolación spline cúbica, se ha supuesto que

$$X''(t_0) = X''(t_k) = 0.$$

### 3.2.1 COMPARACION DE PREDICCIONES CON DATOS REALES

En esta sección, vamos a aplicar la aproximación propuesta en el modelo PCP para predecir la evolución del turismo en Granada durante el último trimestre de 1994.

#### 1. Descripción de los datos

El proceso estocástico considerado en esta aplicación es el grado de ocupación hotelera en Granada a lo largo del año. En principio, disponemos de

observaciones discretas de este proceso al final de cada mes entre los años 1974-1994. Este conjunto de datos ha sido proporcionado por el '*Instituto Nacional de Estadística: Movimiento de viajeros en establecimientos turísticos*' y es representado en la Tabla 1 con doce columnas y veintiuna filas, correspondientes a las variables y las años, respectivamente.

Nuestro objetivo es predecir este proceso para el periodo Septiembre-Diciembre 1994. Por esta razón hemos considerado dos periodos diferentes en el año:  $[T_1, T_2]$  y  $[T_3, T_4]$ , siendo  $T_1 = \text{Enero}$ ,  $T_2 = \text{Agosto}$ ,  $T_3 = \text{Septiembre}$ , and  $T_4 = \text{Diciembre}$ . Para estimar los modelo PCP hemos usado, en cada intervalo, observaciones mensuales de las treinta trayectorias muestrales correspondientes a los años 1974-1993. A modo de ejemplo, en la Figura 1 el grado de ocupación es marcado con puntos y los valores desconocidos entre los nodos son *aproximados* con una curva en trazo continuo.

Las observaciones durante el año 1994 serán usados para medir la exactitud de la predicción proporcionada por los modelos PCP ajustados.

Tabla 1  
Grado de ocupación hotelera en Granada entre 1974-1994

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1974	28.90	27.40	35.30	42.00	38.30	33.50	46.00	58.10	44.60	33.20	25.80	25.10
1975	26.95	30.37	32.90	37.33	39.09	33.58	44.92	53.66	40.73	30.66	23.38	23.92
1976	22.11	23.79	26.34	36.50	30.97	27.01	36.51	46.11	34.89	26.10	21.48	20.35
1977	23.33	31.69	36.37	48.07	35.58	34.14	50.18	59.72	46.28	37.70	32.40	26.42
1978	32.36	35.66	48.33	50.26	50.94	43.11	57.76	69.94	51.89	44.50	35.28	28.73
1979	33.02	34.92	41.46	54.82	45.73	39.93	47.66	59.50	43.05	43.34	33.73	28.89
1980	29.83	32.82	36.14	54.10	43.37	39.39	39.30	58.40	42.56	40.22	29.99	34.42
1981	28.64	28.23	34.02	45.53	46.49	34.19	38.97	57.46	43.99	39.79	29.94	26.96
1982	30.23	31.37	35.72	51.99	43.88	31.77	33.08	55.21	44.19	37.92	29.73	28.38
1983	24.94	31.96	41.73	44.01	39.66	35.42	41.61	56.22	45.91	38.77	29.80	28.57
1984	31.80	31.98	34.31	48.82	43.41	41.97	43.34	62.11	49.03	44.20	33.75	28.98
1985	26.90	32.97	41.15	49.84	45.96	40.49	44.29	56.83	50.06	43.41	30.85	24.71
1986	31.48	33.29	43.66	46.78	50.62	36.76	39.58	54.48	50.58	40.53	31.63	29.53
1987	28.41	32.90	39.42	55.23	49.00	40.40	38.64	52.11	47.07	45.39	33.19	33.82
1988	29.88	37.25	45.69	46.09	45.45	38.29	38.12	52.86	48.38	44.02	35.45	30.32
1989	34.68	38.11	51.57	51.05	51.40	39.57	37.96	54.39	46.39	46.44	37.37	36.00
1990	35.04	40.37	45.71	52.80	47.37	41.12	34.63	46.66	44.68	47.77	41.14	36.12
1991	31.71	35.78	49.39	47.24	45.57	38.63	36.12	47.56	47.02	43.62	35.70	33.74
1992	37.05	42.07	50.14	53.01	50.16	41.17	32.58	46.03	54.32	45.49	32.92	28.49
1993	25.61	37.23	38.13	42.86	43.02	33.54	30.93	41.69	40.16	42.55	29.66	35.79
1994	31.89	42.38	50.16	48.54	50.31	40.22	39.51	46.30	51.79	47.90	31.95	33.05

## 2. Construcción de los modelos PCP aproximados

En primer lugar, hemos aproximado los factores principales muestrales, en cada periodo, mediante MPO, sobre un subespacio de diez funciones trigonométricas, y mediante IBC.

Los valores propios y los porcentajes de varianza explicada por sus cc.pp. asociadas aparecen en la Tabla 2. Obsérvese que, en el segundo periodo, las primeras dos cc.pp. explican más de un 95 % del total de la variabilidad y que con sólo la primera c.p. se explica al menos un 89 %.

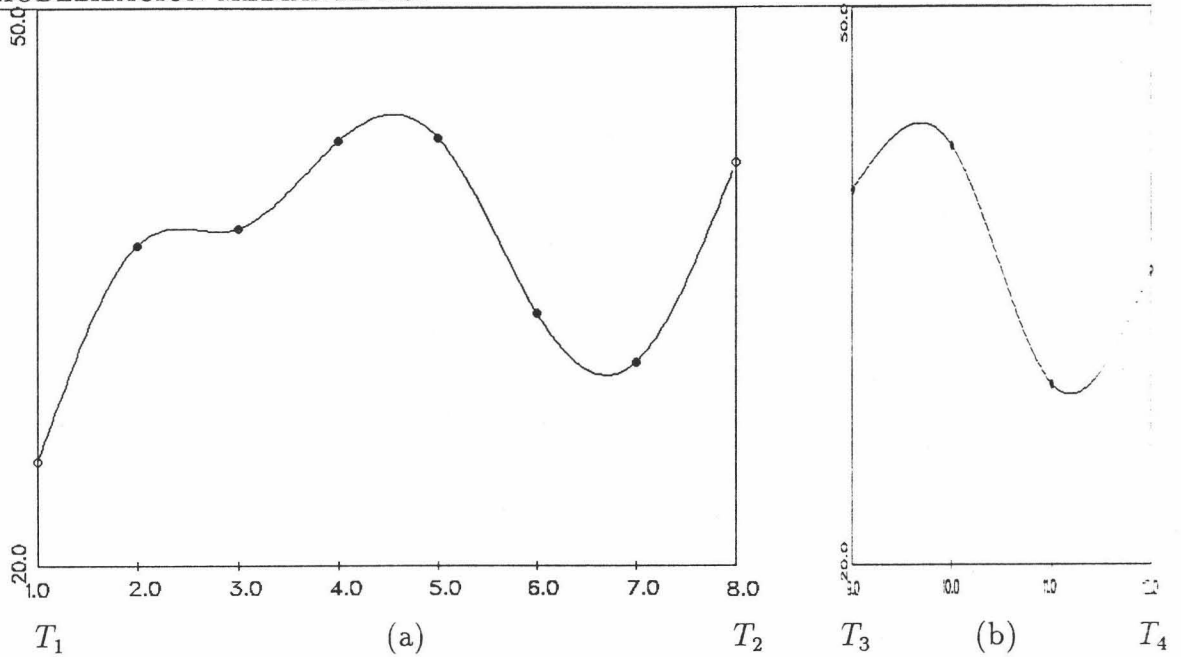


Fig. 1. Grado de ocupación mensual (puntos negros) y curva del grado de ocupación para 1993 durante los periodos: (a) Enero-Agosto; (b) Septiembre-Diciembre

Esto implica que deberíamos construir el modelo PCP con no más de las dos primeras cc.pp. como variables de respuesta.

En segundo lugar, para encontrar los mejores predictores para estas dos cc.pp., hemos estimado las correlaciones lineales entre las cc.pp. en los dos intervalos distintos. Los cuadrados de estos coeficientes de correlacion aparecen en la Tabla 3. Obsérvese que la c.p. con la mayor varianza en el primer periodo está altamente correlada en la primera en el segundo periodo tanto con el método MPO como con el método IBC. Más aún, las restantes correlaciones no son significantes.

Tabla 2

Valores principales y porcentaje de varianza explicada por los dos conjuntos de cc.pp.

Enero-Agosto						
ACP clasico			MPO		IBC	
Nº c.p.	Valores Principales	Varianza Explicada	Valores Principales	Varianza Explicada	Valores Principales	Varianza Explicada
1	117.344	48.66	109.921	55.03	110.602	54.19
2	83.103	34.46	61.685	30.88	64.695	31.70
3	19.246	7.98	15.164	7.59	15.242	7.47
4	8.555	3.55	5.901	2.95	5.9107	2.90
5	4.738	1.97	3.870	1.94	3.9858	1.95
6	4.049	1.68	1.861	0.93	1.8701	0.92
7	2.299	0.95	1.095	0.55	1.1190	0.55
8	1.827	0.76	0.254	0.13	0.6882	0.34
Varianza Total			Varianza Total		Varianza Total	
241.4603			199.7499		204.1126	
Septiembre-Diciembre						
ACP Clasico			OPM		CBI	
Nº c.p.	Valores Principales	Varianza Explicada	Valores Principales	Varianza Explicada	Valores Principales	Varianza Explicada
1	69.636	76.65	61.524	90.77	61.5292	89.34
2	16.307	17.95	4.393	6.48	5.2366	7.60
3	3.274	3.60	1.137	1.68	1.3688	1.99
4	1.636	1.80	0.729	1.07	0.7401	1.07
Varianza Total			Varianza Total		Varianza Total	
90.8529			67.7833		68.8747	



Tabla 3

Cuadrados de las correlaciones lineales entre los dos conjuntos de cc.pp.

		$\hat{\xi}_1$	$\hat{\xi}_2$	$\hat{\xi}_3$	$\hat{\xi}_4$	$\hat{\xi}_5$	$\hat{\xi}_6$	$\hat{\xi}_7$	$\hat{\xi}_8$
$\hat{\eta}_1$	MPO	0.86397	0.00018	0.00291	0.00360	0.00769	0.00507	0.00001	0.02430
	IBC	0.86229	0.00028	0.00327	0.00358	0.00548	0.00638	0.00002	0.02628
	PCA	0.85970	0.00111	0.00419	0.01156	0.00043	0.01119	0.03122	0.00020
$\hat{\eta}_2$	MPO	0.01966	0.08526	0.00118	0.10030	0.00181	0.01740	0.00000	0.01175
	IBC	0.01603	0.12264	0.00003	0.08266	0.00031	0.01371	0.00078	0.00441
	PCA	0.01828	0.2083	0.03508	0.02631	0.00002	0.00038	0.00001	0.01610
$\hat{\eta}_3$	MPO	0.00164	0.19237	0.05350	0.04937	0.01525	0.01136	0.04223	0.06744
	IBC	0.00349	0.15272	0.07279	0.06584	0.01923	0.02179	0.05881	0.06875
	PCA	0.00323	0.04692	0.00456	0.06781	0.00701	0.06896	0.04696	0.00022
$\hat{\eta}_4$	MPO	0.00729	0.00086	0.10221	0.00057	0.09722	0.00848	0.07172	0.00040
	IBC	0.00755	0.00028	0.08544	0.00201	0.09120	0.00774	0.05939	0.00197
	PCA	0.00005	0.00254	0.11472	0.00069	0.04955	0.22241	0.00158	0.00576

De esta forma, hemos identificado el modelo PCP(1;1) dado por

$$\tilde{X}^1(s) = \bar{X}(s) + \tilde{\eta}_1^{-1} \hat{g}_1(s) \quad s \in [T_3, T_4],$$

donde  $\tilde{\eta}_1^{-1}$  ha sido aproximado como

$$\begin{aligned} \text{MPO} \quad \tilde{\eta}_1^{(10)1} &= 0.6954 \hat{\xi}_1^{(10)} \\ \text{IBS} \quad \tilde{\eta}_1^{I1} &= 0.6926 \hat{\xi}_1^I \end{aligned}$$

A pesar de la pequeña correlación existente entre  $\hat{\eta}_2$  y cualesquiera de las c.c.p.p. en el pasado, hemos ajustado también el modelo siguiente PCP(2;1,1) con el fin de compararlo con el modelo PCP(1;1):

$$\tilde{X}^2(s) = \tilde{X}^1(s) + \tilde{\eta}_2^{-1} \hat{g}_2(s), \quad s \in [T_3, T_4]$$

donde la segunda componente principal  $\hat{\eta}_2$  está aproximada como

$$\text{MPO } \tilde{\eta}_2^{(10)1} = -0.2733\hat{\xi}_4^{(10)}$$

$$\text{IBC } \tilde{\eta}_2^{I1} = -0.0996\hat{\xi}_2^I$$

Observemos que la componente principal del pasado con la más alta correlación con  $\hat{\eta}_2$  es la cuarta cuando se usa MPO y la segunda en el caso IBC.

En el siguiente párrafo, vamos a comparar las aproximaciones MPO y IBC para el modelo PCP(1;1). Con este objeto, se han utilizado los errores de predicción en media cuadrática asociados al modelo PCP(1;1), apareciendo calculados en la expresión (3.19) y representados gráficamente en la Figura 2. La predicción del grado de ocupación hotelera en el futuro se representa en la Figura 3 para las aproximaciones MPO y IBC. En la Figura 4 los suavizamientos mediante MPO y IBC son representados para dos años diferentes. En la Figura 5 la proyección ortogonal del grado de ocupación hotelera en el futuro es representada para los mismos años superpuesta con el suavizamiento mediante MPO. De la misma forma, en la Figura 6 se superpone el spline interpolador cubico natural del grado de ocupación en el futuro con su suavizamiento mediante IBC para los mismos años.

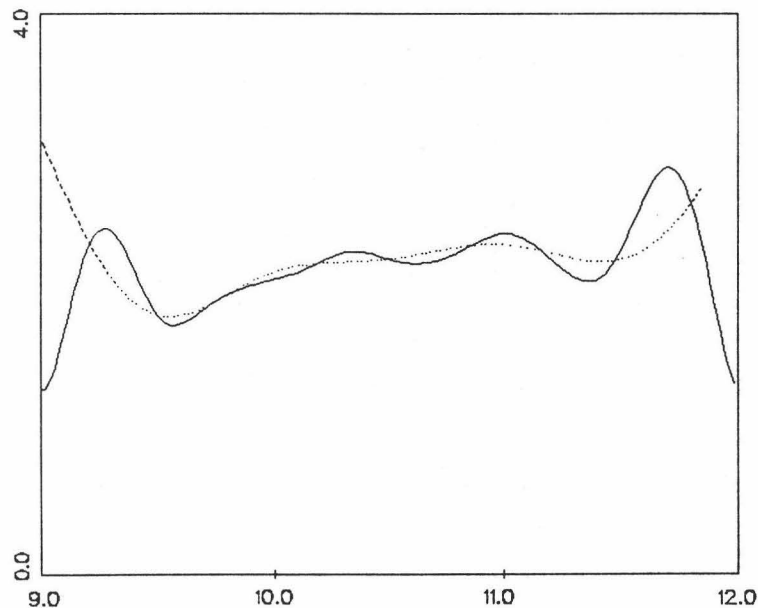


Fig. 2. Error de predicción en media-cuadrática del modelo PCP(1;1) para las aproximaciones MPO (en línea continua) e IBC (en curva punteada)

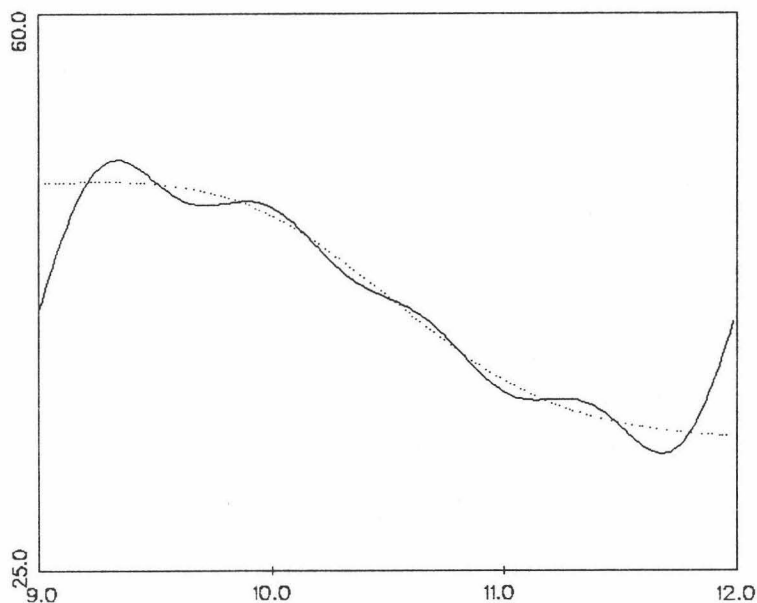


Fig. 3. Predicción del grado de ocupación hotelera de PCP(1;1) para las aproximaciones MPO (en línea continua) e IBC (en línea punteada)

### 3. Comparación con RMCP y discusión

Denotemos con  $\mathbf{X}$  la matriz de datos cuyas columnas, que denotaremos con  $X_j$  ( $j = 1, \dots, 12$ ), representan el grado de ocupación hotelera al final de cada mes durante los años 1974-1993.

Aunque desde un punto de vista teórico los modelos PCP no son directamente comparables con los modelos RMCP para la matriz de datos de respuesta  $\mathbf{X}^1 = (X_9, \dots, X_{12})$  con respecto a la matriz de datos predictores  $\mathbf{X}^2 = (X_1, \dots, X_8)$ , en el siguiente párrafo compararemos las predicciones sobre los nodos de la discretización.

Han sido estimados dos modelos RMCP usando el paquete estadístico BMDP. En la Tabla 2 aparecen los resultados obtenidos del doble-ACP realizado, en concreto, los valores propios muestrales de cada matriz de covarianzas muestrales y los porcentajes de varianza explicada por sus correspondientes cc.pp..

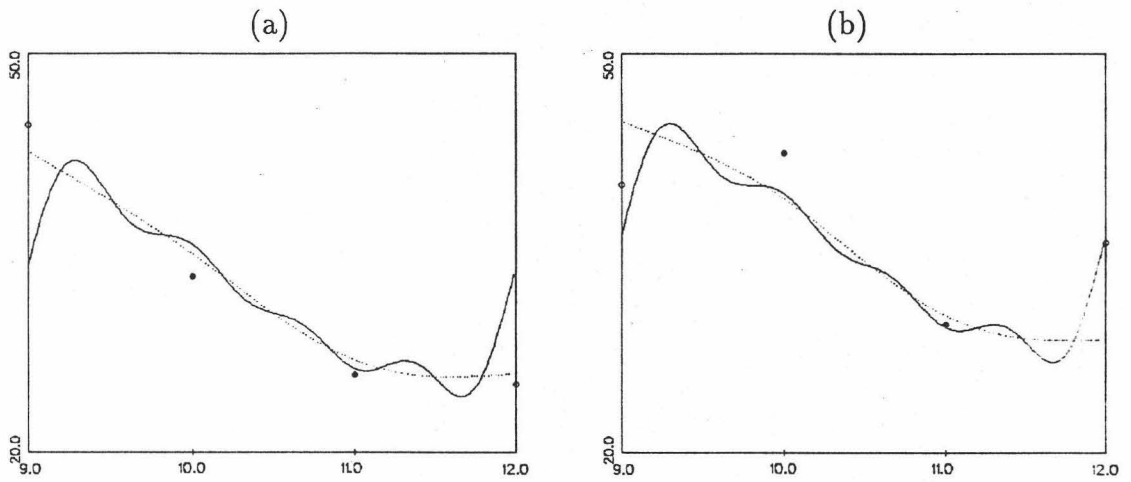


Fig. 4. Diagrama de dispersión y suavizamiento mediante MPO (línea continua) y IBC (línea de puntos) con PCP(1;1) para (a) 1974; (b) 1993

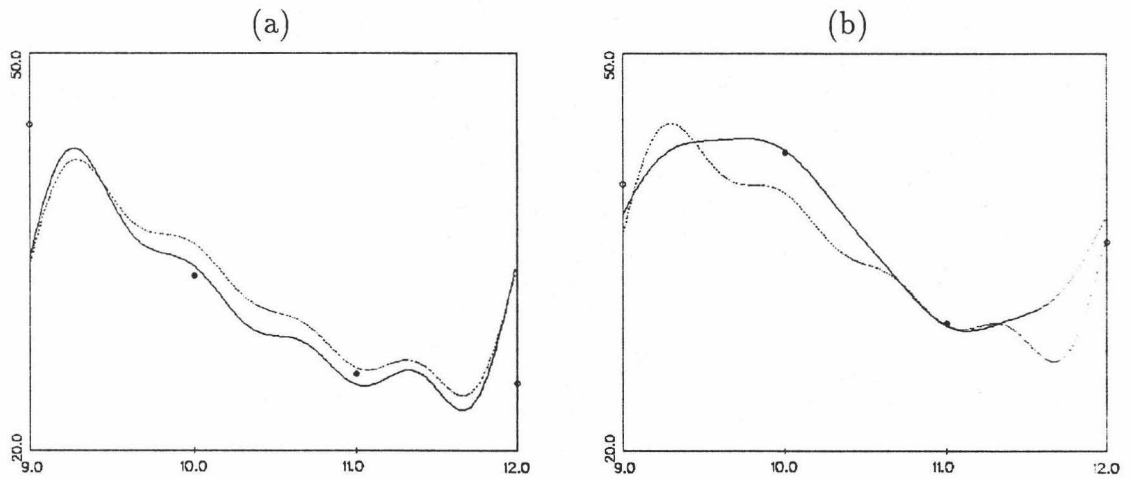


Fig. 5. Diagrama de dispersión, proyección ortogonal (línea continua) y suavizamiento mediante MPO con PCP(1;1) (línea de puntos) para: (a) 1974; (b) 1993

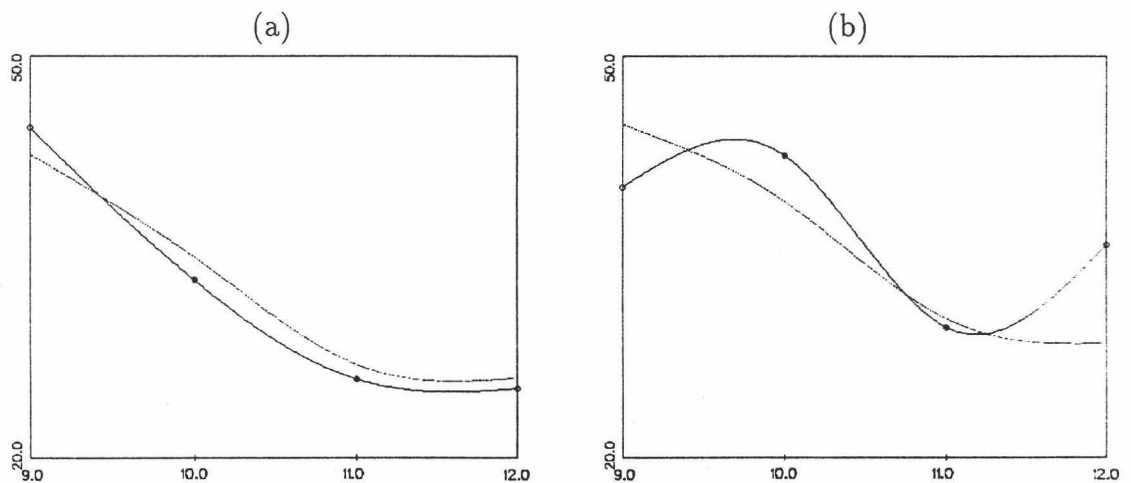


Fig. 6. Diagrama de dispersión, spline cúbico interpolador (línea continua) y su suavizamiento mediante IBC con PCP(1;1) (línea de puntos) para: (a) 1974; (b) 1993

Se observa que, en los dos periodos, el ACP clásico proporciona la mayor varianza total en comparación con las obtenidas mediante los métodos MPO y IBC. La cantidad de varianza explicada por las cc.pp. son también muy diferentes cuando se utiliza RMCP. Por ejemplo, para el ACP clasico la primera c.p. en el futuro explica sólo un 76%, mientras que usando MPO y IBC la proporción es de un 90%. Sin embargo, la varianza acumulada por las dos primeras cc.pp. es similar para las tres alternativas en los dos periodos. Por esta razón, volveremos a usar las primeras dos cc.pp. como variables de respuesta.

Los cuadrados de las correlaciones lineales entre los dos conjuntos de c.c.p.p. aparecen en la Tabla 3. Estos valores son también diferentes a las dadas por las otras aproximaciones (ver, por ejemplo,  $\hat{r}^2(\hat{\eta}_3, \hat{\xi}_2)$  and  $\hat{r}^2(\hat{\eta}_4, \hat{\xi}_6)$ ). Sin embargo las mayores correlaciones son similares a las obtenidas utilizando IBC.

De esta forma, hemos ajustado los siguientes modelos RMCP

$$\text{RMCP(1)} \quad \tilde{X}_j^1 = \bar{X}_j + \tilde{\eta}_1 \hat{\delta}_{j1}$$

$$\text{RMCP(2)} \quad \tilde{X}_j^2 = \tilde{X}_j^1 + \tilde{\eta}_2 \hat{\delta}_{j2}, \quad j = 9, 10, 11, 12$$

después de estimar las primeras dos cc.pp.  $\hat{\eta}_1$  y  $\hat{\eta}_2$  como

$$\begin{aligned} \tilde{\eta}_1 &= 0.7142\hat{\xi}_1 \\ \tilde{\eta}_2 &= 0.2022\hat{\xi}_2 \end{aligned}$$

Por último, hemos obtenido la predicción del grado de ocupación hotelera en Granada en los últimos cuatro meses de 1994 realizando una comparación con los datos reales observados. Las predicciones y el error total en media cuadrática para los modelos ajustados PCP y RMCP aparecen en la Tabla 4. EL error total en media cuadrática para los modelos RMCP ha sido calculado a través de una expresión equivalente a (3.20).

Observemos que las diferencias entre las predicciones proporcionadas por los dos modelos PCP no son significantes para las tres aproximaciones (MPO, IBC and RMCP). Por ello, basaremos nuestra discusión únicamente en el modelo: PCP(1;1).

En la Tabla 4 y Figura 3 observamos que las predicciones dadas por los métodos MPO y IBC son similares excepto en los extremos del intervalo en los que el método MPO toma peores resultados.

Tabla 4  
Grados, real y predicho, de ocupación hotelera

		$\hat{\epsilon}^2$	SEP	OCT	NOV	DIC
Real			51.79	47.90	31.95	33.05
PCP(1,1)	MPO	14.6276	41.37	47.84	36.24	41.37
	IBC	15.8195	49.34	47.32	36.95	33.42
	MPCA	30.9996	49.13	46.81	36.50	33.23
PCP(2,1,1)	MPO	14.1868	41.38	47.86	36.20	41.38
	IBC	15.1777	48.78	47.26	37.08	33.82
	MPCA	27.6028	47.89	46.80	36.59	34.22

Sin embargo, en la Figura 1 se observa que el error en media cuadrática es menor en los extremos cuando se usa MPO. Esto, que puede parecer contradictorio, se explica por el error ha sido calculado utilizando el proceso proyectado con MPO o el proceso interpolación con IBC. Esto significa que MPO predice el proceso aproximado mejor que el método IBC en los extremos. Las Figuras 4 y 5 pueden ser también de ayuda para comprender este resultado. Es decir, el error al final del intervalo, causado por MPO, no es debido al modelo de predicción en componentes principales sino a la propia naturaleza de esta aproximación. Esta conclusión aparece confirmada en el estudio de simulación llevado a cabo en Aguilera et al. (1995).

Por otra parte, en la Tabla 4 puede observarse que las predicciones dadas por las aproximaciones MPO y IBC son al menos tan buenas como las dadas por RMCP. Sin embargo, la principal ventaja de los métodos MPO y IBC es la de predecir y suavizar el proceso entre los nodos de observación.

### §3.3 TRATAMIENTO DE NODOS DESIGUALMENTE ESPACIADOS

En esta sección, nos centraremos en el caso de que para un cierto conjunto finito (fijo) de  $k + 1$  nodos (no necesariamente igualmente espaciados)

$$\pi_k \stackrel{\text{def}}{=} \{t_0 < \dots < t_k\},$$

la información muestral disponible para cada una de las curvas muestrales venga descrita, tan sólo, por sus valores sobre dicho conjunto (fijo) de nodos.

es decir,  $\{X_w(t_j) : w = 1, \dots, N; j = 0, 1, \dots, k\}$ ,

En la anterior descripción del método de interpolación se ha hecho uso de que el conjunto de los splines cúbicos

$$S_3(\pi_k) = \left\{ S \in C^2[t_0, t_k] : S|_{[t_{j-1}, t_j]} \in \mathcal{P}_3 \quad j \in \{1, \dots, k\} \right\}$$

tiene estructura de espacio vectorial de dimensión  $k + 3$  (Prenter, 1975). Concretamente, la consideración de una base facilita la utilización de los splines, permitiendo su utilización a través de sus coordenadas, y hace que, desde un punto de vista computacional, la obtención de los *estimadores aproximados* venga precedida por operaciones matriciales.

En nuestro caso, la base que se ha considerado en el espacio,  $S_3(\pi_k)$ , recibe el nombre de base de los B-splines. A continuación mostraremos cómo está definida dicha base.

**Definición 3.3.1** *Añadiendo seis nodos a los de la partición  $\pi_k$*

$$t_{-3} < t_{-2} < t_{-1} < t_0 < \dots < t_k < t_{k+1} < t_{k+2} < t_{k+3}$$

se definen los B-splines de orden 1 como sigue

$$B_{i,1}(t) = \begin{cases} 0 & t < t_{i-2} \\ 1 & t_{i-2} \leq t < t_{i-1} \\ 0 & t \geq t_{i-1} \end{cases} \quad (i = -1, 0, \dots, k+4)$$

y por recursividad, se definen los B-splines de orden  $n$  a partir de los de orden  $n - 1$ , de forma que, para cada  $n \in \{2, 3, 4\}$  se tiene

$$B_{i,n}(t) = \frac{t - t_{i-2}}{t_{i+n-3} - t_{i-2}} B_{i,n-1}(t) + \frac{t_{i+n-2} - t}{t_{i+n-2} - t_{i-1}} B_{i+1,n-1}(t) \\ i \in \{-1, 0, 1, \dots, k+5-n\}$$

Se denominan B-splines cúbicos (de orden 4) a la base de  $S_3(\pi_k)$  dada por  $\{B_i \stackrel{\text{def}}{=} B_{i,4} : i = -1, \dots, k+1\}$ .

Las ventajas, que ofrece la anterior definición recursiva, están relacionadas con óptimas propiedades numéricas, tales como la estabilidad (Boor, 1974), útiles de cara a una implementación computacional.

El imponer que los nodos estén igualmente espaciados, es decir, sean de la forma  $\{t_i = t_0 + ih : i = 0, \dots, k\}$  (con  $h > 0$ ), permite una simplificación considerable en el diseño computacional, y una economía en lo que a volumen de cálculo se refiere, tanto en la obtención de la base, como en su posterior utilización dentro, como es nuestro caso, del método IBC. Bajos esta condición, se obtienen en Aguilera (1993) las expresiones computacionales con la base de los B-splines correspondientes al método IBC.

A continuación, intentaremos mostrar cómo se resolvió el problema computacional bajo la condición de que los nodos no sean necesariamente igualmente espaciados.

### EVALUACION DE LOS B-SPLINES

Para la implementación computacional de la evaluación de los B-splines se ha considerado la anterior definición recursiva, pero expresada de forma más conveniente para tal objeto.

Comenzaremos analizando los soportes de los B-splines de cada uno de los órdenes considerados.

- $n = 1$  e  $i \in \{-1, k + 4\}$

$$B_{i,1} \leftarrow [t_{i-2}, t_{i-1})$$

- $n = 2$  e  $i \in \{-1, k + 3\}$

$$B_{i,2} \leftarrow \left. \begin{array}{l} B_{i,1} \leftarrow [t_{i-2}, t_{i-1}) \\ B_{i+1,1} \leftarrow [t_{i-1}, t_i) \end{array} \right\} \leftarrow [t_{i-2}, t_i)$$

- $n = 3$  e  $i \in \{-1, k + 2\}$

$$B_{i,3} \leftarrow \left. \begin{array}{l} B_{i,2} \leftarrow [t_{i-2}, t_i) \\ B_{i+1,2} \leftarrow [t_{i-1}, t_{i+1}) \end{array} \right\} \leftarrow [t_{i-2}, t_{i+1})$$

- $n = 4$  e  $i \in \{-1, k + 1\}$

$$B_{i,4} \leftarrow \left. \begin{array}{l} B_{i,3} \leftarrow [t_{i-2}, t_{i+1}) \\ B_{i+1,3} \leftarrow [t_{i-1}, t_{i+2}) \end{array} \right\} \leftarrow [t_{i-2}, t_{i+2})$$



Por tanto, si deseamos evaluar los elementos de la base de los B-splines sobre valores del intervalo  $[t_0, t_k]$ , tendremos que considerar el recubrimiento de intervalos  $\{[t_{i-1}, t_i) : i = 1, \dots, k + 1\}$ .

De lo anterior, dado  $s_o \in [t_0, t_k]$  existirá  $i \in \{1, \dots, k + 1\}$  tal que

$$s_o \in [t_{i-1}, t_i)$$

En estas condiciones, los B-splines, de cada uno de los órdenes, cuyo soporte contine al punto  $s_o$ , son lo siguientes:

$$\begin{array}{ccccccc}
 & & & & & & B_{i-2,4} \\
 & & & & & B_{i-1,3} & \longrightarrow & B_{i-1,4} \\
 & & & B_{i,2} & \longrightarrow & B_{i,3} & \longrightarrow & B_{i,4} \\
 B_{i+1,1} & \longrightarrow & B_{i+1,2} & \longrightarrow & B_{i+1,3} & \longrightarrow & B_{i+1,4}
 \end{array}$$

Obsérvese que en el anterior diagrama de flujo, cuando  $s_o = t_k$  ( $i = k + 1$ ) el elemento  $B_{i+1,4}$  no aparecería por no existir  $B_{k+2,4}$ .

En cualquier caso, en aras a la evaluación de un B-spline de orden 4 sobre el punto  $s_o$ ,  $B_{j,4}(s_o)$ , podemos hacer la siguiente disquisición de casos

- Si  $j$  no pertenece a  $\{i - 2, i - 1, i, i + 1\}$  entonces  $B_{j,4}(s_o) = 0$

•

$$B_{i+1,1}(s_o) \longrightarrow B_{i,2}(s_o) \longrightarrow B_{i-1,3}(s_o) \longrightarrow B_{i-2,4}(s_o)$$

•

$$\begin{array}{ccccccc}
 & & & & & B_{i-1,3}(s_o) & \longrightarrow & B_{i-1,4}(s_o) \\
 & & & B_{i,2}(s_o) & \longrightarrow & B_{i,3}(s_o) & & \\
 B_{i+1,1}(s_o) & \longrightarrow & B_{i+1,2}(s_o) & & & & & 
 \end{array}$$

•

$$\begin{array}{ccccccc}
 & & & B_{i,2}(s_o) & \longrightarrow & B_{i,3}(s_o) & \longrightarrow & B_{i,4}(s_o) \\
 B_{i+1,1}(s_o) & \longrightarrow & B_{i+1,2}(s_o) & \longrightarrow & B_{i+1,3}(s_o) & & & 
 \end{array}$$

•

$$B_{i+1,1}(s_o) \longrightarrow B_{i+1,2}(s_o) \longrightarrow B_{i+1,3}(s_o) \longrightarrow B_{i+1,4}(s_o)$$

Teniendo en cuenta cómo se disponen los B-splines de los diferentes órdenes, que habría que obtener previamente hasta calcular

$$B_{j,4}(s_o) \quad \text{con } j \in \{-1, \dots, k+1\} \quad s_o \in [t_0, t_k]$$

podemos seguir la siguiente estrategia:

(A) *Inicio*

Obtención del índice  $i \in \{1, \dots, k+1\}$  tal que  $s_o \in [t_{i-1}, t_i]$ . En estas condiciones, pueden distinguirse dos posibles situaciones:

- Si  $j \in \{i-2, i-1, i, i+1\}$  entonces continuar con el punto (B)
- en caso contrario  $B_{j,4}(s_o) = 0 \rightarrow$  (FIN).

(B) *Evaluación de los B-splines de orden 2*

Considérense dos valores  $v[0]$  y  $v[1]$  inicializados a 0, y obtener

- si  $j < i+1$  entonces  $v[0] \stackrel{\text{def}}{=} B_{i,2} = \frac{t_i - s_o}{t_i - t_{i-1}}$
- si  $j > i-2$  entonces  $v[1] \stackrel{\text{def}}{=} B_{i+1,2} = \frac{s_o - t_{i-1}}{t_i - t_{i-1}}$

(C) *Inicialización de un índice, que denotamos con  $l$*

$$l = \begin{cases} 1 & \text{Si } j \in \{i-2, i-1\} \\ 0 & \text{Si } j \in \{i, i+1\} \end{cases}$$

(D) *Evaluación de los B-splines de orden 3*

- *Asignación a  $v[l]$  del valor  $B_{i,3}(s_o)$*   
Si  $j \in \{i-1, i\}$  entonces

$$v[l] \stackrel{\text{def}}{=} B_{i,3}(s_o) = \frac{s_o - t_{i-2}}{t_i - t_{i-2}} v[0] + \frac{t_{i+1} - s_o}{t_{i+1} - t_{i-1}} v[1]$$

en caso contrario  $v[l] \stackrel{\text{def}}{=} 0$

- Asignación a  $v[1-l]$  del valor de  $B_{i-1,3}(s_o)$  (cuando  $l = 1$ ), o bien  $B_{i+1,3}(s_o)$  (cuando  $l = 0$ ). Concretamente, sabemos que

$$B_{i-1,3}(s_o) = \frac{t_i - s_o}{t_i - t_{i-2}} B_{i,2}(s_o) \quad (l = 1)$$

$$B_{i+1,3}(s_o) = \frac{s_o - t_{i-1}}{t_{i+1} - t_{i-1}} B_{i+1,2}(s_o) \quad (l = 0)$$

obsérvese cómo ambas expresiones podrían ser descritas como sigue

$$v[1-l] = \frac{s_o - t_{i-1+l}}{t_{i+1-3l} - t_{i-1+l}} v[1-l]$$

(E) *Evaluación del B-spline de orden 4*

$$B_{j,4}(s_o) = \frac{s_o - t_{i-2-l}}{t_{i+1-l} - t_{i-2-l}} v[0] + \frac{t_{i+2-l} - s_o}{t_{i+2-l} - t_{i-1-l}} v[1]$$

Obsérvese que los puntos anteriores definen un algoritmo para la evaluación de los B-splines de orden cuatro. Este algoritmo, es de fácil implementación computacional, así mismo, al evitar la utilización de la recursividad, permite que a la hora de implementarlo con un lenguaje de programación no sea necesaria la utilización de los recursos de la recursividad.

### EVALUACION DE PRODUCTOS ESCALARES DE B-SPLINES

Otro de los problemas, en este caso, asociado a la puesta en práctica del método IBC fué la obtención de los elementos de la matriz **P**. Como se ha mostrado en la sección anterior, esta matriz es de dimensión  $(k+3) \times (k+3)$  y posee por elementos a los dados por las integrales de la forma:

$$\int_{t_0}^{t_k} B_i(s) B_j(s) ds \quad \forall i, j \in \{-1, \dots, k+1\}$$

En principio, consideraremos sin pérdida de generalidad que  $i \leq j$ . Por otro lado, haciendo uso de que los soporte de los B-splines  $B_i$  y  $B_j$  son respectivamente

$$[t_{i-2}, t_{i+2}] \quad [t_{j-2}, t_{j+2}]$$

puede garantizarse que si  $i + 2 \leq j - 2$ , o equivalentemente  $i + 4 \leq j$  se obtendría

$$B_i B_j(t) = 0 \quad \forall t \in [t_0, t_k]$$

Con lo cuál, para cada  $i \in -1, \dots, k + 1$  las únicas integrales a considerar serían las siguientes  $\forall j \in \{i, i + 1, i + 2, i + 3\} \cap \{-1, \dots, k + 1\}$

$$\int_{t_0}^{t_k} B_i(s) B_j(s) ds = \int_{t_{j-2}}^{t_{i+2}} B_i(s) B_j(s) ds$$

En estas condiciones, disponemos de un algoritmo con el que podemos evaluar los B-splines sobre cualquier punto de un intervalo, y así mismo por otro lado, los B-splines son polinomios de grado tres entre cada dos nodos consecutivos. Por ello, se obtó por la utilización de una fórmula de cuadratura de tipo interpolatorio. Concretamente, en nuestro caso la idea era aplicar una fórmula de tipo interpolatorio (que fuese exacta al menos para polinomios de grado menor o igual que seis) entre cada dos nodos consecutivos.

Concretamente, como se demuestra en uno de los apéndices de esta tesis, la fórmula de cuadratura gaussiana con cuatro puntos, cuyo grado de exactitud alcanza a los polinomios de grado menor o igual que siete, puede obtenerse de la forma siguiente:

**Proposición 3.3.1** *La fórmula de cuadratura gaussiana con cuatro nodos sobre un intervalo real  $[a, b]$  viene dada por*

$$\int_a^b f(x) dx \cong \frac{14}{15} (b - a) \left\{ \frac{f(x_0) + f(x_3)}{H_2} + \frac{f(x_1) + f(x_2)}{H_1} \right\}$$

donde estamos denotando con

$$H_k \stackrel{\text{def}}{=} (10s_k - 1) (-28s_k + 21.6) \quad \forall k = 1, 2$$

siendo

$$s_1 = \frac{15 - 2\sqrt{30}}{35} \quad s_2 = \frac{15 + 2\sqrt{30}}{35}$$

$$v_0 = -\sqrt{s_2} \quad v_1 = -\sqrt{s_1} \quad v_2 = \sqrt{s_1} \quad v_3 = \sqrt{s_2}$$

$$x_j = \frac{a + b}{2} + \frac{b - a}{2} v_j \quad \forall j = 0, 1, 2, 3$$

## Apéndice A

# OBTENCION DE UNA FORMULA DE CUADRATURA

### §A.1 PLANTEAMIENTO DEL PROBLEMA

Pretendemos encontrar un algoritmo de eficiente implementación computacional, que nos permita obtener la fórmula de cuadratura de Gauss para cualquier número de nodos, a elegir en un intervalo real arbitrario. Es decir, dados un intervalo real  $[a, b]$  y un número  $n + 1$ , de nodos, intentaremos obtener una metodología para calcular los elementos

- $a \leq x_0 < \dots < x_n \leq b$
- $\{\alpha_j : j = 0, \dots, n\} \subseteq \mathbb{R}$

con los que la correspondiente fórmula de cuadratura

$$\int_a^b f(x) dx = \sum_{j=0}^n \alpha_j f(x_j) + E(f) \quad (\text{A.1})$$

alcance un máximo grado de exactitud polinomial.

### §A.2 SOBRE EL GRADO DE EXACTITUD DE UNA F. C.

En esta sección se enumeran los resultados que nos permitirán caracterizar el grado de exactitud polinomial de una fórmula de cuadratura con  $n + 1$  nodos;

así como obtener el máximo grado de exactitud posible que se puede alcanzar.

**Teorema A.2.1** *Una fórmula de cuadratura (A.1) es exacta para todo polinomio de  $\wp_{n+q}$ , si y solamente si, se verifican las siguientes condiciones:*

1. *La fórmula (A.1) es de tipo interpolatorio.*
2. *Cuando  $q \geq 1$*

$$\int_a^b \Pi(x) x^k dx = 0 \quad \forall k = 0, \dots, q-1 \quad (\text{A.1})$$

donde  $\Pi$  es el polinomio definido por

$$\Pi(x) \stackrel{\text{def}}{=} \prod_{j=0}^n (x - x_j) \quad \forall x \in \mathbb{R}$$

Obsérvese que la condición (A.1) del teorema anterior podría ser reformulada como sigue: cuando  $q \geq 1$

$$\Pi \in (\wp_{q-1})^{\perp L_2[a,b]} \quad (\text{A.2})$$

**Teorema A.2.2** *No existe ninguna fórmula de cuadratura de la forma (A.1) que sea exacta para polinomios de grado superior o igual a  $2n + 2$ .*

□

Demostraciones de cualesquiera de los teoremas anteriores pueden encontrarse en cualquier libro de Análisis Numérico que trate el problema de la integración numérica.

Una consecuencia inmediata del teorema anterior (A.2.2) es que el máximo grado de exactitud al que podemos aspirar, para una f.c. con  $n + 1$  nodos, es de  $2n + 1$ .

### §A.3 UN PROBLEMA EQUIVALENTE

Haciendo uso del teorema (A.2.1) y del comentario hecho en (A.2), obtenemos un primer procedimiento para obtener fórmulas de cuadratura (A.1) de grado  $n + q$  con  $n + 1$  nodos en un intervalo  $[a, b]$ . Los pasos a seguir serían los siguientes:

1. Los nodos serán  $n + 1$  valores en el intervalo  $[a, b]$

$$a \leq x_0 < \dots < x_n \leq b$$

de forma que el polinomio asociado  $\Pi$  cumpla (A.2).

2. Una vez obtenidos los nodos, como la f. c. ha de ser de tipo interpolatorio, la aplicación de dicha fórmula a una función  $f$  sería equivalente a considerar, con dichos nodos, su polinomio de interpolación,  $p_f \in \wp_n$ , obteniéndose la aproximación de la integral como sigue:

$$\int_a^b f(x) dx \cong \int_a^b p_f(x) dx$$

En principio, el polinomio  $p_f \in \wp_n$  está determinado de forma única por la condición

$$f(x_j) = p_f(x_j) \quad \forall j = 0, \dots, n$$

y puede ser obtenido de la forma siguiente:

$$p_f(x) = \sum_{j=0}^n f(x_j) L_j(x) \quad \forall x \in \mathbb{R}$$

Donde con  $L_j$  estamos denotando al  $j$ -ésimo polinomio de la base de Lagrange de  $\wp_n$  asociada a los nodos anteriores. Es decir:

$$L_j(x) \stackrel{\text{def}}{=} \prod_{h \in \sigma_j} \frac{x - x_h}{x_j - x_h} \quad \forall x \in \mathbb{R}$$

siendo el conjunto de índices  $\sigma_j \stackrel{\text{def}}{=} \{0, \dots, n\} - \{j\}$

De esta forma, se obtendría que

$$\int_a^b p_f(x) dx = \sum_{j=0}^n f(x_j) \alpha_j$$

considerando  $\forall j = 0, \dots, n$

$$\alpha_j = \int_a^b L_j(x) dx \quad (\text{A.1})$$

Desafortunadamente, el procedimiento anterior de obtención de una fórmula de cuadratura, con  $n+1$  nodos sobre un intervalo  $[a, b]$  y con grado de exactitud polinomial de  $n+q$ , es de difícil implementación computacional.

Nuestro propósito es analizar cada uno de los pasos del procedimiento anterior, con el objeto de conseguir un grado mayor de simplificación computacional.

En primer lugar, tomemos como marco de referencia el intervalo  $[-1, 1]$ , para ello consideramos el difeomorfismo creciente siguiente:

$$\phi : [-1, 1] \mapsto [a, b]$$

$$\phi(t) \stackrel{\text{def}}{=} c + tr \quad \forall t \in [-1, 1]$$

donde estamos notando con:

$$c \stackrel{\text{def}}{=} \frac{a+b}{2} \quad r \stackrel{\text{def}}{=} \frac{b-a}{2}$$

Teniendo en cuenta lo anterior, podrían obtenerse los nodos anteriores de la forma siguiente:

$$x_j = \phi(v_j) \quad \forall j = 0, \dots, n$$

donde  $-1 \leq v_0 < v_1 < \dots < v_n \leq 1$

Con lo cuál,  $\forall x = \phi(t) \in [a, b]$  se tiene

$$x - x_j = r(t - v_j) \quad \forall j = 0, \dots, n$$

En esa línea, la función polinómica  $\Pi$ , a través de  $\phi$ , adoptaría la forma siguiente:



$$\Pi \circ \phi(t) = r^{n+1} F(t) \quad \forall t \in [-1, 1]$$

donde

$$F(t) \stackrel{\text{def}}{=} \prod_{j=0}^n (t - v_j) \quad \forall t \in [-1, 1]$$

Obsérvese que otra reformulación equivalente a la condición (A.2) sería imponer que  $\forall k = 0, \dots, q-1$

$$J_k \stackrel{\text{def}}{=} \langle \Pi, p_k \rangle_{L_2[a,b]} = 0 \quad (\text{A.2})$$

donde el sistema  $\{p_k : k = 0, \dots, q-1\}$  es la base del subespacio vectorial  $\wp_{q-1}$  dada por

$$p_k(x) \stackrel{\text{def}}{=} (x - c)^k$$

A continuación, expresaremos las integrales  $J_k$  como integrales sobre el intervalo  $[-1, 1]$ . A través del difeomorfismo  $\phi$  se tiene  $\forall k = 0, \dots, q-1$

$$J_k = \int_{-1}^1 \Pi(\phi(t)) p_k(\phi(t)) \phi'(t) dt = r^{n+2+k} \int_{-1}^1 F(t) t^k dt$$

En definitiva, la condición (A.2) sobre los nodos  $\{x_j : j = 0, \dots, n\}$  es equivalente a la condición sobre el conjunto  $\{v_j : j = 0, \dots, n\}$  dada por

$$\int_{-1}^1 F(t) t^k dt = 0 \quad \forall k = 0, \dots, q-1 \quad (\text{A.3})$$

Continuando con el propósito inicial, a partir del  $j$ -ésimo polinomio de Lagrange, anteriormente denotado con  $L_j$ , podría considerarse:

$$l_j(t) \stackrel{\text{def}}{=} L_j \circ \phi(t) = \prod_{h \in \sigma(j)} \frac{t - v_h}{v_j - v_h}$$

Con ello, la obtención de los coeficientes (A.1) vendría dada  $\forall j = 0, \dots, n$  por

$$\alpha_j = \int_{-1}^1 L_j(\phi(t)) \phi'(t) dt = r \int_{-1}^1 l_j(t) dt$$

Con lo razonado hasta ahora podemos enunciar el siguiente resultado equivalente a lo expuesto en el teorema (A.2.1).

**Teorema A.3.1** Una fórmula de cuadratura (A.1) es exacta para todo polinomio de  $\wp_{n+q}$ , si y solamente si, existe un polinomio  $G \in \wp_{n+1}$  cumpliendo:

- Todas las raíces de  $G(t)$  son reales, distintas y pertenecientes al intervalo  $[-1, 1]$

$$-1 \leq v_0 < \dots < v_n \leq 1$$

- Si  $q \geq 1$  entonces  $G \in \wp_{q-1}^\perp L_2[-1, 1]$

De forma que en caso afirmativo  $\forall j = 0, \dots, n$  se verifica:

$$x_j = \phi(v_j)$$

$$\alpha_j = r \int_{-1}^1 l_j(t) dt$$

□

El denotar con  $G$ , al polinomio cuyo papel en la deducción anterior lo ha jugado  $F$ , se hace con el objeto de recalcar que el coeficiente líder (el coef. líder de  $F$  es 1) no influye en lo enunciado en el teorema anterior. En concreto, la relación existente entre el anterior polinomio  $F$  y  $G$ , es la multiplicación por un escalar.

Con lo visto anteriormente, el problema de obtener una f. c. sobre un intervalo arbitrario con  $n + 1$  nodos, es equivalente al problema de obtener un polinomio ortogonal en  $L_2[-1, 1]$  con unas determinadas propiedades. La idea de este trabajo está basada en esta equivalencia.

## §A.4 POLINOMIOS DE LEGENDRE

Lo razonado anteriormente establece el marco donde vamos a tratar nuestro problema. En concreto, la idea que se ha seguido para abordarlo ha sido utilizar la base de los polinomios de Legendre como herramienta para resolver nuestro problema.

En esta sección mostraremos la sucesión de polinomios conocidos como polinomios de Legendre. En particular, nos centraremos en su obtención y en algunas propiedades que mostrarán su utilidad posteriormente.

#### A.4.1 Obtención mediante la fórmula de Rodrigues

Para cada  $m \in \mathbb{N}$ , se define el  $m$ -ésimo polinomio de Legendre de la forma siguiente:

$$P_m(t) \stackrel{\text{def}}{=} \frac{1}{2^m m!} \frac{d^m}{dt^m} (t^2 - 1)^m \quad \forall t \in \mathbb{R}$$

**Ejemplo A.4.1** Algunos polinomios de Legendre son:

$$\begin{aligned} P_0(t) &= 1 & P_1(t) &= t \\ P_2(t) &= \frac{1}{2}(3t^2 - 1) & P_3(t) &= \frac{1}{2}(5t^3 - 3t) \end{aligned}$$

#### A.4.2 Propiedades de los polinomios de Legendre

1. La sucesión de polinomios de Legendre  $\{P_m : m \in \mathbb{N}\}$  forma un sistema ortogonal de  $L^2[-1, 1]$ .
2.  $\int_{-1}^1 P_m(t)^2 dt = \frac{2}{2m+1}$
3. El grado de  $P_m$  es  $m$ .
4. El conjunto  $\{P_0, \dots, P_m\}$  forma una base del espacio  $\wp_m$ , para cualquier  $m \in \mathbb{N}$ .
5. Cuando  $m \geq 1$ , el polinomio  $P_m$  tiene  $m$ -raíces reales en el intervalo abierto  $(-1, 1)$ .

### §A.5 OBTENCIÓN DE LA F. C. GAUSSIANA CON $n + 1$ NODOS

El objeto de esta sección es aplicar lo razonado anteriormente a la obtención de una f. c. gaussiana con  $n + 1$  nodos.

En primer lugar, por lo visto en el teorema (A.2.2), el máximo grado de exactitud, que podemos pretender alcanzar con  $n + 1$  nodos, es de  $2n + 1$  ( $q = n + 1$ ).

Así mismo, teniendo en cuenta el teorema (A.3.1), consideremos el polinomio de Legendre  $P_{n+1} \in \wp_{n+1}$ . Dicho polinomio sabemos que cumple, entre otras, las siguiente propiedades:

- Posee  $n + 1$  raíces reales, distintas y pertenecientes al intervalo  $(-1, 1)$ . Es decir, si notamos a sus raíces con

$$-1 < v_0 < \dots < v_n < 1$$

se obtendría

$$P_{n+1}(t) = \beta \prod_{j=0}^n (t - v_j) \quad \forall t \in \mathbb{R}$$

donde  $\beta$  es el coeficiente líder de  $P_{n+1}$ .

- En principio, puesto que los polinomios de Legendre son una base ortogonal de  $L_2[-1, 1]$ , se cumpliría  $\forall k = 0, \dots, n$

$$\int_{-1}^1 P_{n+1}(t) P_k(t) dt = 0$$

Además, sabemos que el sistema  $\{P_0, \dots, P_n\} \subseteq \wp_n$  forma una base ortogonal de  $\wp_n$ . Con lo que se tendrá

$$P_{n+1} \in \wp_{n+1} \cap (\wp_n)^{\perp L_2[-1,1]}$$

En definitiva, el polinomio  $P_{n+1}$  resulta ser el polinomio que se propone en el teorema (A.3.1) con cuyas raíces se obtendrá la f. c. en cuestión.

Además, obsérvese que en este caso es fácil demostrar que

$$\dim(\wp_{n+1} \cap (\wp_n)^{\perp L_2[-1,1]}) = 1$$

con lo cuál

$$\wp_{n+1} \cap (\wp_n)^{\perp L_2[-1,1]} = \text{Lin}\{P_{n+1}\}$$

Y en definitiva, en el caso de buscar una f. c. gaussiana, cualquier polinomio no nulo del subespacio lineal de dimensión uno anteriormente citado, podría ser objeto de consideración para utilizar el teorema (A.3.1). Por ello, consideraremos a partir de ahora a  $G$  un polinomio

$$G \in \wp_{n+1} \cap (\wp_n)^{\perp L_2[-1,1]}$$

Aunque desde un punto de vista teórico parecería resuelto nuestro problema, sin embargo, la aplicación del mencionado teorema (A.3.1) nos llevaría a enfrentarnos con los dos *problemas* computacionales siguientes:

**Problema 1.** Obtención de las raíces del polinomio  $G$ . Para ello utilizaremos alguno de los métodos numéricos de resolución de ecuaciones algebraicas.

**Problema 2.** Una vez obtenidas las raíces  $\{v_j : j = 0, \dots, n\}$ , tendríamos que evaluar las integrales siguientes:

$$I_j \stackrel{\text{def}}{=} \int_{-1}^1 l_j(t) dt \quad \forall j = 0, \dots, n$$

Para este segundo problema intentaremos obtener expresiones alternativas, a las que se obtendría mediante una evaluación directa, que permitan mayores simplificaciones en los cálculos.

#### A.5.1 Estudio de las integrales $I$

En principio, consideraremos el conjunto de polinomios  $\{\psi_0, \dots, \psi_n\} \subseteq \wp_n$  definidos  $\forall j = 0, \dots, n$

$$\psi_j(t) \stackrel{\text{def}}{=} \prod_{h \in \sigma(j)} (t - v_h) \quad \forall t \in \mathbb{R}$$

Obsérvese que podríamos expresar  $\forall j = 0, \dots, n$

$$l_j(t) = \frac{\psi_j(t)}{\psi_j(v_j)} \quad \forall t \in \mathbb{R}$$

Por otro lado, recordemos que  $F(t) = \prod_{j=0}^n (t - v_j)$  cumpliéndose además que

$$F \in \wp_{n+1} \cap (\wp_n)^\perp L_2[-1,1] \quad (\text{A.1})$$

**Lema A.5.1** *El conjunto  $\{\psi_0, \dots, \psi_n\}$  constituye una base  $L_2[-1,1]$ -ortogonal de  $\wp_n$ .*

*Demostración*

Dados  $j, i = 0, \dots, n$ , arbitrarios pero fijos con  $j \neq i$ , puede observarse que haciendo uso de la expresión (A.1) se cumple

$$\int_{-1}^1 \psi_i(t) \psi_j(t) dt = \int_{-1}^1 F(t) Q(t) dt = 0$$

ya que

$$Q(t) \stackrel{\text{def}}{=} \prod_{h \in \sigma_{i,j}} (t - v_h) \in \wp_{n-1}$$

siendo  $\sigma_{i,j} \stackrel{\text{def}}{=} : \{0, \dots, n\} - \{i, j\}$

Por tanto, el conjunto anterior formado por  $n + 1$  polinomios constituye un sistema ortogonal, y por tanto linealmente independiente, en el subespacio vectorial  $\wp_n$  de dimensión  $n + 1$ . De donde se concluye que dicho sistema es además una base de  $\wp_n$ .

□

**Corolario A.5.1** *El conjunto  $\{l_0, \dots, l_n\}$  constituye una base  $L_2[-1, 1]$ -ortogonal de  $\wp_n$ .*

*Demostración*

Obsérvese que  $\forall j = 0, \dots, n$  se tiene

$$l_j = \frac{1}{\psi_j(v_j)} \psi_j$$

donde  $\psi_j(v_j) \neq 0$

□

**Lema A.5.2** *Para cada  $j = 0, \dots, n$  se cumple*

$$\int_{-1}^1 l_j(t)^2 dt = \int_{-1}^1 l_j(t) dt$$

*Demostración*

Consideremos  $1 \in \wp_n$  (polinomio constante), entonces por el corolario (A.5.1) se sabe que existen  $\{a_i : i = 0, \dots, n\} \subseteq \mathbb{R}$  tales que

$$1 = \sum_{i=0}^n a_i l_i$$

En particular, obsérvese que se cumpliría  $\forall h = 0, \dots, n$

$$1 = \sum_{i=0}^n a_i l_i(v_h) = a_h$$

por lo que se tendría

$$1 = \sum_{i=0}^n l_i(t) \quad \forall t \in \mathbb{R}$$

Teniendo en cuenta esto, para cada  $j = 0, \dots, n$  podemos expresar

$$l_j(t) = \sum_{i=0}^n l_j(t) l_i(t) \quad \forall t \in \mathbb{R}$$

de donde se obtendría, haciendo uso del corolario anterior, lo siguiente:

$$\int_{-1}^1 l_j(t) dt = \sum_{i=0}^n \int_{-1}^1 l_j(t) l_i(t) dt = \int_{-1}^1 l_j(t)^2 dt$$

□

**Lema A.5.3**

$$F'(t) = \sum_{j=0}^n \psi_j(t) = \sum_{j=0}^n \psi_j(v_j) l_j(t) \quad \forall t \in \mathbb{R}$$

□

**Teorema A.5.1** Para cada  $i = 0, \dots, n$  se tiene

$$\int_{-1}^1 l_i(t) dt = \frac{1}{\psi_i(v_i)^2} \left[ \frac{F(1)^2}{1-v_i} + \frac{F(-1)^2}{1+v_i} \right]$$

*Demostración*

Dado  $i = 0, \dots, n$ , arbitrario pero fijo, estudiemos la integral siguiente:

$$K \stackrel{\text{def}}{=} \int_{-1}^1 F'(t) l_i(t) dt$$

Por un lado, teniendo en cuenta el lema anterior y el corolario (A.5.1) se obtiene

$$K = \sum_{j=0}^n \psi_j(v_j) \int_{-1}^1 l_j(t) l_i(t) dt = \psi_i(v_i) \int_{-1}^1 l_i^2(t) dt$$

de donde utilizando el lema (A.5.2) se concluye

$$K = \psi_i(v_i) \int_{-1}^1 l_i(t) dt \quad (\text{A.2})$$

Por otro lado, integrando por partes en  $K$  y utilizando que  $l_i' \in \wp_{n-1}$ , se obtiene

$$K = F(t)l_i(t)]_{-1}^1 - \int_{-1}^1 F(t)l_i'(t) dt = F(t)l_i(t)]_{-1}^1$$

Así mismo, utilizando la definiciones de  $F$  y  $l_i$ , se obtiene

$$K = \frac{1}{\psi_i(v_i)} F(t)\psi_i(t)]_{-1}^1 = \frac{1}{\psi_i(v_i)} \left. \frac{F^2(t)}{t - v_i} \right]_{-1}^1 \quad (\text{A.3})$$

Uniendo los resultados (A.2) y (A.3) se obtiene la tesis del teorema.

Obsérvese que la evaluación directa del término  $\psi_i(v_i)$  sería posiblemente fuente de errores de aproximación, puesto que significaría realizar el producto de diferencias de, posiblemente, aproximaciones de las raíces. Para resolver este problema considérese el siguiente resultado.

**Corolario A.5.2** *Para cada  $i = 0, \dots, n$  se cumple*

$$\psi_i(v_i) = \prod_{h \in \sigma_i} (v_i - v_h) = F'(v_i)$$

*Demostración*

Es una consecuencia inmediata del lema (A.5.3). □

A continuación, reformularemos el teorema anterior en términos del polinomio  $G$ , cuya única diferencia frente a  $F$  es su coeficiente líder; además de integrar lo enunciado en el lema anterior.

**Teorema A.5.2** *Para cada  $i = 0, \dots, n$  se cumple*

$$I_i = \int_{-1}^1 l_i(t) dt = \frac{1}{G'(v_i)^2} \left[ \frac{G(1)^2}{1 - v_i} + \frac{G(-1)^2}{1 + v_i} \right]$$

*Donde  $G$  es un polinomio proporcional al polinomio de Legendre cuyo grado,  $n + 1$ , y por tanto el número de sus raíces*

$$-1 < v_0 < \dots < v_n < 1$$

*coincide con el número de nodos con el que se pretende construir la f. c. gaussiana.*



## §A.6 FORMULA DE CUADRATURA GAUSSIANA CON CUATRO NODOS

Como aplicación de lo anterior, obtendremos la f. c. gaussiana con cuatro nodos en un intervalo  $[a, b]$ , cuyo grado de exactitud será de  $2 * 3 + 1 = 7$ .

Para ello consideremos el cuarto polinomio de Legendre

$$P_4(t) = \frac{1}{8} (35t^4 - 30t^2 + 3)$$

y consideremos el polinomio

$$G(t) = 35t^4 - 30t^2 + 3$$

1. Calculemos las raíces de  $G(t)$ .

Haciendo uso del cambio  $s = t^2$  resolvamos la ecuación

$$35s^2 - 30s + 3 = 0$$

de donde se obtendría las soluciones siguientes:

$$s_1 = \frac{15 - 2\sqrt{30}}{35} \quad s_2 = \frac{15 + 2\sqrt{30}}{35}$$

las cuales, a su vez proporcionan las raíces del polinomio  $G$ , que ordenadas en orden creciente son:

$$\begin{aligned} v_0 &= -\sqrt{s_2} \cong -0.8611363 \\ v_1 &= -\sqrt{s_1} \cong -0.339981 \\ v_2 &= \sqrt{s_1} \cong 0.339981 \\ v_3 &= \sqrt{s_2} \cong 0.8611363 \end{aligned} \tag{A.1}$$

2. Evaluación de las integrales  $I_i$

Haciendo uso del teorema (A.5.2) se obtiene:

$$I_i \stackrel{\text{def}}{=} \int_{-1}^1 l_i(t) dt = \frac{1}{G'(v_i)^2} \left[ \frac{G(1)^2}{1 - v_i} + \frac{G(-1)^2}{1 + v_i} \right] =$$

$$= \frac{G(1)^2}{G'(v_i)^2} \frac{1 + v_i + 1 - v_i}{1 - v_i^2} = \frac{2G(1)^2}{G'(v_i)^2 (1 - v_i^2)}$$

Calculemos cada uno de los elementos de la expresión anterior

En primer lugar,  $G(1) = 8$

Por otro lado, se tiene

$$G'(t) = (140t^3 - 60t) = 20t(7t^2 - 3)$$

de donde  $G'(v_i) = 20v_i(7v_i^2 - 3)$

En definitiva, obtendríamos lo siguiente:

$$I_i = \frac{8}{25v_i^2(7v_i^2 - 3)^2(1 - v_i^2)}$$

Puede así afirmarse que las integrales  $I_i$  dependen sólo de los valores  $v_i^2$ . Por esta razón, y con el fin de simplificar la notación, denotaremos con

$$w_i \stackrel{\text{def}}{=} v_i^2 \quad \forall i = 0, 1, 2, 3$$

Con lo que se obtendría, para cada  $i = 0, 1, 2, 3$ , lo siguiente:

$$I_i = \frac{8}{25w_i(7w_i - 3)^2(1 - w_i)}$$

donde

$$w_0 = w_3 = s_2 \quad w_1 = w_2 = s_1$$

A continuación, analizaremos la expresión siguiente:

$$\Delta_i \stackrel{\text{def}}{=} (7w_i - 3)^2(1 - w_i)$$

En principio, operando en dicha expresión se obtiene

$$\Delta_i = -49w_i^3 + 91w_i^2 - 51w_i + 9$$

Para intentar eliminar el término independiente de la anterior expresión, consideremos la ecuación que cumplen los valores  $w_i$

$$35w_i^2 - 30w_i + 3 = 0$$

de donde se obtendría

$$3 (35w_i^2 - 30w_i + 3) = 105w_i^2 - 90w_i + 9 = 0$$

Restamos lo obtenido anteriormente a la expresión de  $\Delta_i$ , con lo que se obtiene

$$\Delta_i = w_i (-49w_i^2 - 14w_i + 39)$$

A continuación intentaremos disminuir el grado en  $w_i$  de la expresión anterior que aparece entre paréntesis. Para ello, volviendo a utilizar la ecuación que cumplen los valores  $w_i$  se obtiene

$$w_i^2 = \frac{30w_i - 3}{35} = \frac{3}{35} (10w_i - 1) \quad (\text{A.2})$$

de donde

$$-49w_i^2 - 14w_i + 39 = -\frac{21}{5} (10w_i - 1) - 14w_i + 39 = -56w_i + 43.2$$

De todo lo anteriormente razonado, se obtiene para  $\Delta_i$  la siguiente expresión

$$\Delta_i = w_i (-56w_i + 43.2)$$

con la cuál, y utilizando además la expresión (A.2) se obtiene

$$I_i = \frac{8}{25 w_i^2 (-56w_i + 43.2)} =$$

$$\frac{8}{2 * 25 \frac{3}{35} (10w_i - 1) (-28w_i + 21.6)} =$$

$$\frac{28}{15} \frac{1}{(10w_i - 1) (-28w_i + 21.6)}$$

**Proposición A.6.1** *La fórmula de cuadratura gaussiana con cuatro nodos sobre el intervalo real  $[a, b]$  viene dada por*

$$\int_a^b f(x) dx \cong \frac{14}{15} (b-a) \left\{ \frac{f(x_0) + f(x_3)}{H_2} + \frac{f(x_1) + f(x_2)}{H_1} \right\}$$

donde estamos denotando con

$$H_k \stackrel{\text{def}}{=} (10s_k - 1) (-28s_k + 21.6) \quad \forall k = 1, 2$$

siendo

$$s_1 = \frac{15 - 2\sqrt{30}}{35} \quad s_2 = \frac{15 + 2\sqrt{30}}{35}$$

$$x_j = \phi(v_j) \quad \forall j = 0, 1, 2, 3$$

con los valores  $v_j$  calculados en (A.1)

*Demostración*

$$\begin{aligned} \int_a^b f(x) dx &\cong r \sum_{j=0}^n f(x_j) I_j = \\ r \left\{ [f(x_0) + f(x_3)] \int_{-1}^1 l_0(t) dt + [f(x_1) + f(x_2)] \int_{-1}^1 l_1(t) dt \right\} &= \\ = \frac{14}{15} (b-a) \left\{ \frac{f(x_0) + f(x_3)}{H_2} + \frac{f(x_1) + f(x_2)}{H_1} \right\} & \end{aligned}$$

□

# Bibliografía

- [1] Abellanas L. y A. Galindo (1987). *Espacios de Hilbert (geometría, operadores, espectros)*. Eudema, Madrid.
- [2] Anderson T.W. (1984). *An introduction to multivariate statistics analysis*. Wiley. New York.
- [3] Aguilera A.M., M.J. Valderrama and M.J. Del Moral (1992). "Un método para la aproximación de estimadores en ACP. Aplicación al proceso de Ornstein-Uhlenbeck". *Rev. Soc. Chilena Estadística*, 9 (2), 57-77.
- [4] Aguilera A.M. (1993). "Métodos de aproximación de estimadores en el ACP de un proceso estocástico". Tesis Doctoral, Universidad de Granada.
- [5] Aguilera A.M. and M.J. Valderrama (1993a). "Principal component analysis of a stochastic process for discrete data: interpolation by B-splines". *Proc. 41th Session ISI*, Vol. I, 11-12.
- [6] Aguilera A.M., M.J. Del Moral and M.A. Piñar (1993b). "On the empirical behaviour of a stochastic process". *Proc. 6th Internat. Symposium on ASMDA*, Vol. 1, 5-16.
- [7] Aguilera A.M., F.A. Ocaña and M.J. Valderrama (1994). "A computational algorithm for PCA of random processes". *Proc. COMPSTAT*, Software Descriptions, 39-40.
- [8] Aguilera A.M., R. Gutiérrez, F.A. Ocaña and M.J. Valderrama (1995a). "Computational approaches to estimation in the principal component analysis of a stochastic process". *Appl. Stoch. Models Data Anal.*, 11 (4), en prensa.
- [9] Aguilera A.M., F.A. Ocaña and M.J. Valderrama (1996). "Análisis en Componentes Principales de un Proceso Estocástico con Funciones Muestrales Escalonadas". *Qüestio*, Vol. 20, en prensa.
- [10] Aguilera A.M., F.A. Ocaña and M.J. Valderrama (1995b). "Predicción dinámica de un proceso estocástico mediante análisis en componentes principales". *Actas de la I Reunión de Trabajo en Predicción Dinámica*, 197-215.

- [43] Prenter, P.M. (1975). *Splines and variational methods*. Wiley, New York.
- [44] Rao C.R. (1964). "The use and interpretation of principal component analysis in applied research". *Sankhya, A*, 26, 329-358.
- [45] Ramsay J.O.(1982). "When the data are functions". *Psychometrika*, Vo. 47, 4.
- [46] Ramsay J.O. y C.J. Dalzell (1991). "Some tools for functional data analysis". *Journal of the Royal Statistical Society, serie B*, 53 (3), 539-572.
- [47] Ramsay J.O. y X. Wang (1995). "A Functional Data Analysis of the Pinch Force of Human Fingers". *Appl. Statist.* 44, No. 1, 17-30.
- [48] Rice J.A. and B.W. Silverman (1991). "Estimating the mean and covariance structure nonparametrically when the data are curves". *Journal of the Royal Statistical Society, serie B*, 53 (1), 233-243.
- [49] Riesz F. y B. Sz-Nagy (1990). *Lecons d'Analyse Fonctionnelle*. Gauthier-Villars.
- [50] Saporta G. (1981). *Méthodes exploratoires d'analyse de données temporelles*. Cahiers du Bureau Universitaire de Recherche Opérationnelle, 37-38.
- [51] Shorack G.R. y J.A. Wellner (1986). *Empirical Processes with Applications to Statistics*. Wiley.
- [52] Todorovic P. (1992). *An introduction to stochastic processes and their applications*. Springer-Verlag, New York.
- [53] Wahba G. (1990). *Spline models for observational data*. SIAM, Philadelphia, Pennsylvania.
- [54] Watkins D.S. (1993). "Some Perspectives on the Eigenvalue Problem". *SIAM Rev.*, 35 (3), 430-471.
- [55] Wong E. (1971). *Stochastic processes in information and dynamical systems*. McGraw-Hill, New York.

ciadas a v.a. hilbertianas, viendo la representación proporcionada por el ACP como un caso particular caracterizado por determinadas propiedades de optimalidad. El final de este capítulo está dedicado al problema de la estimación del ACP. Concretamente, se utiliza la definición de los estimadores del ACP proporcionada por Deville (1973b), y se enuncian algunos resultados allí demostrados, así como otros, referentes al comportamiento asintótico de dichos estimadores, demostrados en Dauxois et al. (1982).

La elección de las v.a. hilbertianas como marco teórico se ha hecho en base a que: (i) por un lado, es el marco en donde la generalización del ACP multivariante es más intuitiva, y además se obtienen como casos particulares el ACP para vectores, funciones, campos escalares, campos vectoriales, etc.; (ii) permite resaltar con claridad la influencia la estructura geométrica considerada en el *espacio de datos* sobre el correspondiente ACP, a diferencia, por ejemplo, del marco proporcionado por los procesos, más intuitivo pero en el que esta característica queda relegada frente a la consideración de colección continua de variables aleatorias. En cualquier caso, esta elección está avalada así mismo por algunos de los trabajos enmarcados en la modelización de datos funcionales como por ejemplo, Deville (1973b), Besse (1991), Besse and Cardot (1994)

Así mismo, indicaremos que a lo largo de la memoria los espacios de Hilbert que se considerarán, siempre que no se indique lo contrario, son todos separables.

Por su parte, el trabajo que aquí se expone es fruto del estudio realizado en esa línea en la modelización de datos de naturaleza funcional. Concretamente, este trabajo se concibió como una continuación del trabajo de de tesis doctoral de Aguilera (1993) en el que se abordó, principalmente, el problema de la aproximación al ACP a partir de datos discretos. Entre los problemas que quedaron abiertos se encontraba la consideración de una estructura geométrica, diferente de la proporcionada por las funciones de cuadrado integrables de Lebesgue, que describiese el carácter funcional de los datos.

Entre los aspectos más interesantes de cara a la aplicación del ACP podemos señalar que, por un lado, permite explorar la estructura interna de los datos, y por otra parte, es utilizado como una alternativa en la reducción de la dimensión del problema. Estos dos usos pueden aparecer contrapuestos, ya que, como a veces ocurre, la idea de reducción de la dimensión no tiene por qué ir acompañada con la obtención de las componentes que nos permitan entender mejor el fenómeno. Concretamente, en los trabajos de Deville (1977), Besse et al. (1986b), Aguilera et al. (1996) pueden verse ejemplos de una apli-

- [11] Aguilera A.M., F.A. Ocaña and M.J. Valderrama (1995c). "ACP de procesos estocásticos en tiempo continuo frente a cambios de métrica en el espacio de trayectorias". *Actas de la SEIO*, 3-5.
- [12] Atkinson L.V. y P.J. Harley (1983). *An introduction to numerical methods with Pascal*. Addison-Wesley.
- [13] Baker C.T.H. (1977). *The Numerical Treatment of Integral Equations*. Oxford University Press.
- [14] Bensoussan A. (1971). *Filtrage optimal des systèmes linéaires*. Dunod, Paris.
- [15] Besse Ph. (1987). "Choix de la métrique pour l'acp d'événements discrets". *Statistique et Analyse des Données* 12 (3), 1-16.
- [16] Besse Ph. (1988). "Spline functions and optimal metric in linear principal component analysis". En: *Component and correspondence analysis. Dimension reduction by functional approximation*. Wiley, New York, 81-101.
- [17] Besse Ph., H. Caussinus, L. Ferre and J. Fine (1986a). "Some guidelines for principal components analysis". *Proc. COMPSTAT*, Springer-Verlag, Heidelberg, 23-30.
- [18] Besse Ph. and J.O. Ramsay (1986b). "Principal component analysis of sample functions". *Psychometrika*, Vol. 51, No. 2, pp. 285-311.
- [19] Besse Ph., H. Caussinus, L. Ferre and J. Fine (1988). "Principal components analysis and optimization of graphical displays". *Statistics*, 2, 301-312.
- [20] Besse Ph. (1991). "Approximation spline de l'analyse en composantes principales d'une variable aléatoire hilbertienne". *Annales de la Faculté des Sciences de Toulouse*, 12, 329-346.
- [21] Besse P. and H. Cardot (1994). "Approximation spline de la prévision d'un processus autorégressif hilbertien d'ordre 1". *Publication du Laboratoire de Statistique et Probabilités de Toulouse*, 17.
- [22] Bosq D. (1991). *Modelization, non-parametric estimation and prediction for continuous time processes*. Nato, Asi series.
- [23] Castro P.E., W.H. Lawton and E.A. Sylvestre (1986). "Principal modes of variation for processes with continuous sample curves". *Technometrics*, 28 (4), 329-337.
- [24] Church A. (1966). "Analysis of Data When the Response Is a Curve". *Technometrics* 8, 229-246.
- [25] Cuadras C.M. (1981). *Metodos de Analisis Multivariante*. Eunibar, Barcelona.



- [26] Dauxois J., J. Fine-Fontan y A. Pousse (1979). "Echantillonnage en segmentation. Etude de la convergence". *Statistique et Analyse des Données*, 3, 45-53.
- [27] Dauxois J., A. Pousse and Y. Romain (1982). "Asymptotic theory for the principal component analysis of a vector random function: some applications to statistical inference". *J. Multivar. Anal.*, 12, 136-154.
- [28] Davis P.J. (1975). *Interpolation and approximation*. Dover Publications, New York.
- [29] De Boor C. (1978). *A practical guide to splines*. Springer-Verlag, New York.
- [30] Deville J.C. (1972). "Processus punctuels et calendrier de constitution de familles". Note interne de l'INSEE.
- [31] Deville J.C. (1973a). "Analyse harmonique et calendrier de constitution des familles". *Proc. 39th Session of ISI*, 309-324.
- [32] Deville J.C. (1973b). "Estimation of the eigenvalues and of the eigenvectors of a covariance operator". Note Interne de l'INSEE.
- [33] Deville J.C. (1973c). "L'analyse harmonique dans le cas de données discrètes". Note interne de l'INSEE.
- [34] Deville J.C. (1974). "Méthodes statistiques et numériques de l'analyse harmonique". *Annales de l'INSEE*, 15, 3-101.
- [35] Deville J.C. (1977). "Analyse harmonique du calendrier de constitution des familles en France: disparités sociales et évolution de 1920 a 1960". *Population*, 1, 17-63.
- [36] Deville J.C. (1978). "Analyse et prevision des series chronologiques multiples non stationnaires". *Statist. Anal. Données*, 3, 19-29.
- [37] Fukunaga K. (1990). *Introduction to Statistical Pattern Recognition*. Academic Press, San Diego.
- [38] Gutiérrez R., J.C. Ruiz, y M.J. Valderrama (1992). "On the numerical expansion of a second order stochastic process". *Applied Stochastic Models and Data Analysis*, 8 (2), 67-77.
- [39] Jackson J.E. (1991). *A user's guide to principal components*. Wiley, New York.
- [40] McGarty T.P. (1974). *Stochastic systems and state estimation*. Wiley, New York.
- [41] Jolliffe I.T. (1986). *Principal Component Analysis*. Springer-Verlag, New York.
- [42] Pardoux C. (1989). "Apport de l'Analyse Factorielle a l'Etude d'un Processus". *Revue Statistique Appliquée*, XXXVII 4, 41-60.